An Overview on Data Hiding Techniques

Prof. Vaibhav Desai
HOD, BCA Dept.
SDJ International College
Surat, Gujarat

ABSTRACT :

For us, information or data is a really valuable resource. As a result, information security becomes even more important. The exchange of information, medium we use to convey data does not guarantee any protection. As a result, other techniques of data security are being considered. The importance of information concealment cannot be overstated. today's role It included techniques for encrypting data. information in such a way that it is illegible by anyone unintentional user. This paper examines the methods that are available for data concealment, and how may these be used?

KEYWORDS : Watermarking, Cryptography, Steganography, Data Hiding

INTRODUCTION

Any company or individual requires a lot of data or information. We don't like it when our communication is overheard since it could be used against us. The same may be said for any organization's or individual's data. To avoid tampering, data communication between two prospective parties must take place in a safe manner. During any information exchange, there are two categories of threats. The unexpected user who overhears this chat has two options: tamper with the data to modify its original meaning, or listen to the message with the aim of decoding it and using it to his or her advantage. Both of these assaults compromised the message's secrecy and integrity. It's a difficult undertaking to provide intended access while avoiding unintentional access. Information has been kept hidden for a long time. People used hidden pictures or invisible ink to communicate secret information in the past [1].

NEED OF DATA HIDING

The relevance of data masking strategies stems from the fact that the medium via which the information is sent is unreliable, or in other words, the media is unsecure. As a result, several techniques are required to make it impossible for an unintended user to extract information from the message. There are a few reasons why data is kept hidden which are as follows :

1. Personal and confidential information
2. Sensitive information
3. Trade secrets and confidential information
4. To prevent data misuse
5. Data loss due to unintentional harm, human error, or accidental deletion
6. For monetary and extortion objectives
7. To conceal evidence of criminal activity
8. Last but not least, have some fun with it.

CHARACTERISTICS OF DATA HIDING

Any strategy for concealing information must have the following characteristics:

1. **Capacity** : The amount of information that can be buried in cover media is referred to as capacity [1]. The quantity of information that can be hidden is limited by the fact that hidden information should not totally alter the original message in order to prevent drawing inadvertent attention to it.
2. **Security** : Data should be secured using the information concealing approach so that only the intended user has access to it.

In other terms, it relates to an unauthorized user's incapacity to discover hidden information. This is critical in order to safeguard the confidentiality and sensitivity of the data being delivered [1, 2].

3. **Robustness** : It refers to the amount of data that can be hidden without having any negative consequences and without deleting the data [1].

4. **Perceptibility** : The data concealing method should conceal data in such a way that the original cover signal and the concealed data signal are indistinguishable to the human eye.

TECHNIQUES OF INFORMATION HIDING

Watermarking, cryptography, and steganography are three prevalent data concealment techniques.

Watermarking : A watermark is an identifiable image or pattern imprinted on paper that serves as proof of authenticity [3, 4]. When viewed in transmitted light, the watermark appears as varying hues of lightness/darkness. Watermarks are commonly used to add protection to banknotes, passports, postage stamps, and other security documents [4]. In the digital realm, digital watermarking is an extension of this principle.

Today, there is so much material on the internet that we are forced to use procedures to preserve digital media ownership. Digital information piracy is rampant, whether it's photographs, text, audio, or video. These are fairly simple to make and distribute. As a result, determining who owns the material becomes critical. Digital watermarking is a solution to the long-standing issue of digital data copyright [3]. A digital watermark is a type of marking that is hidden within digital data like audio or image data. It can be extracted or identified later to make data assertions. This information could be about the author, the copyright, or the image itself [3, 5]. The digital watermark is preserved during transmission/transformation, allowing us to secure our digital ownership rights. Digital watermarks are only visible under particular conditions, for as after applying an algorithm, and are invisible otherwise. It's pointless to employ a digital watermark if it alters the carrier signal to the point of being perceptible.

The basic purpose of a watermarking system is to achieve resilience, which means that the watermark should be impossible to remove without messing with the underlying data.

Digital watermarking is a form of passive security. It only tags the data, not degrades it or restricts access to it [5].

Source tracking is one application of digital watermarking. At each point of distribution, a watermark is placed in a digital signal. If a copy of the work is later discovered, the watermark may be extracted from the copy, and the source of the dissemination can be determined. This method is said to have been used to track down the source of illegally copied movies. Another application is in broadcast monitoring, where watermarked video from international agencies is frequently seen on television news.

1. **Cryptography** : The words crypt and graphy mean "hidden or secret" and "writing," respectively. The word is borrowed from the Greek language. The art of converting data into an unreadable format known as cypher text is known as cryptography. The communication is deciphered or decrypted by the recipient on the other side. Data secrecy, data integrity, authentication, and nonrepudiation are all provided by cryptography. Confidentiality refers to the restriction or limitation of access to particular types of information. Integrity refers to the preservation and assurance of the accuracy of data being supplied, i.e., information that has not been modified, deleted, or otherwise altered. Authentication verifies that the source and receiver of the information are the same person. The capacity to ensure that the sender or receiver cannot deny the legitimacy of their signature on the transmission information that they created [6] is known as non-repudiation. Encryption is identical with cryptography in today's world. Plain text refers to the unencrypted data, while cypher text refers to the encrypted data. There are three steps to cryptography.

1. Encryption – converting plain text into an unreadable format. Cipher text is the result of this process.

2. Message transmission - this entails sending the encryption text to the intended recipient.

3. Decryption — on the receiving end, the encrypted text is decrypted to reveal the plain text.

Strictly speaking, there are two types of cryptography: symmetric key cryptography and asymmetric key cryptography.

1. Symmetric key cryptography - this refers to encryption systems in which the sender and receiver both have access to the same key. This form of encryption is used by several encryption algorithms such as AES, DES, and RC5.

Key that is symmetric Plain text, encryption algorithm, secret key, cypher text, and decryption algorithm are the five components of cryptography. The encryption algorithm uses a secret key to perform various actions on plain text. One of the communicating parties chooses the secret key, which is independent of plain text. This produces cypher text as a result. The decryption algorithm takes the cypher text and the secret key as input and outputs plaintext.

A notable disadvantage of the symmetric key cypher is that it requires each pair of communication parties to share the secret key, as well as the key itself to be shared in a secure channel. Any unwanted user with access to the secret key poses a risk of decrypting the text.

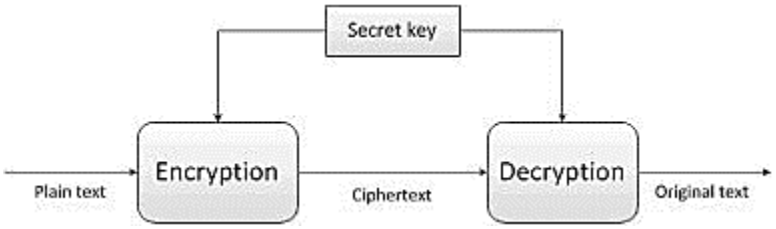

Figure 1 : Flow of Symmetric key Cryptography

2. Public key cryptography, also known as asymmetric key cryptography. It makes use of two keys: a public and a private key. The public key can be freely transmitted, but the private key that pairs with it must be kept secret. Encryption is performed using the public key. After then, the encryption text is sent to the recipient. The receiver decrypts the plain text using a secret key and a decryption algorithm.

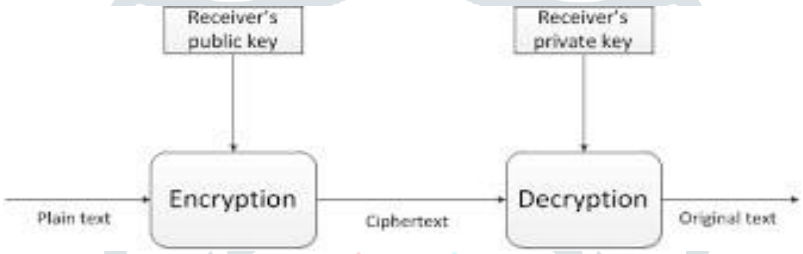

Figure 2 : Flow of Asymmetric key Cryptography

Digital signatures can also employ public key cryptography. Digital signatures can be linked to the content of the message being signed indefinitely. The secret key is used to sign the contents, and the accompanying public key is used to verify the signature's legitimacy.

3. **Steganography** : Steganography is the process of concealing or hiding a message, file, or image within another message, file, or image. Steganography comes from the Greek term steganos, which meaning "covered writing" or "concealed writing" [7]. In other terms, it is the art and science of communicating in a method that conceals the communication's existence. The idea is to hide messages inside other innocent communications in such a way that the opponent is unable to identify the presence of a second message [3]. Steganography is more concerned with high levels of security and capacity. Even minor alterations to the stego medium can alter its meaning.

Steganography hides critical information in cover media such as photos, audio, and video sent over the internet.

There are four steps to steganography:

1. Choosing the cover medium that will be used to conceal the data.
2. The camouflaged secret message or information to be conveyed through the cover media.
3. A function that will be used to hide data in the cover media and extract the hidden data, as well as its inverse.
4. An optional key or password for data authentication or hiding and unhiding [2].

The chosen cover should be done with great care. Because steganography works by replacing redundant data with the secret message, the cover chosen should have enough redundant information that can be utilized to hide the data.

Different types used in steganography techniques:

1. Text Steganography
2. Image Steganography
3. Audio Steganography
4. Video Steganography
5. IP Steganography

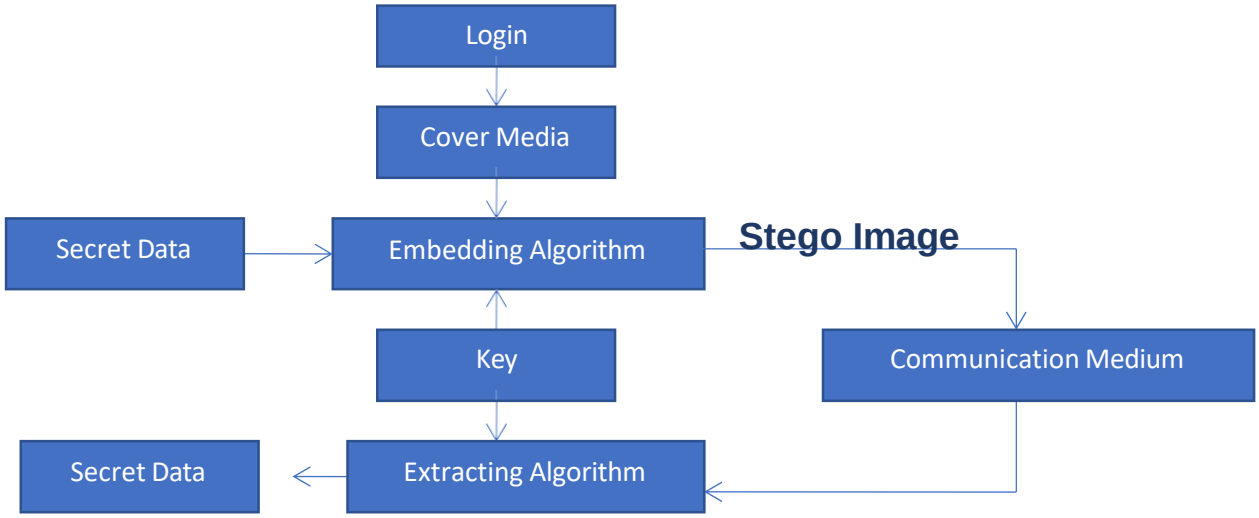

Figure 3 : Flow of Steganography

STEGANOGRAPHY VS. CRYPTOGRAPHY

Cryptography means "secret writing," while steganography means "cover writing" [3]. Although steganography and cryptography are sometimes conflated, there are significant differences between the two. The former hides the information and sends it to the network through a cover. Unintentional users will have a tough time determining whether or not there is any hidden information encoded. The key feature of steganography is that the cover should contain enough redundant information that even after the message has been embedded, it is difficult to discern the message just by glancing at it. Cryptography, on the other hand, encrypts a communication so that it either becomes unreadable or has its original meaning completely altered.

Steganography does not change the structure of the secret message, whereas cryptography does. The former prevents the existence of the communication from being discovered, whereas the latter stops an unauthorized user from learning the contents of a communication.

STEGANOGRAPHY + CRYPTOGRAPHY

Both strategies can be combined to create an additional layer of security. The message can be first encrypted to a cypher text using cryptography. This encrypted text can then be steganographically embedded in a cover material. The three aims of data hiding will be met with this integrated approach: security, capacity, and robustness.

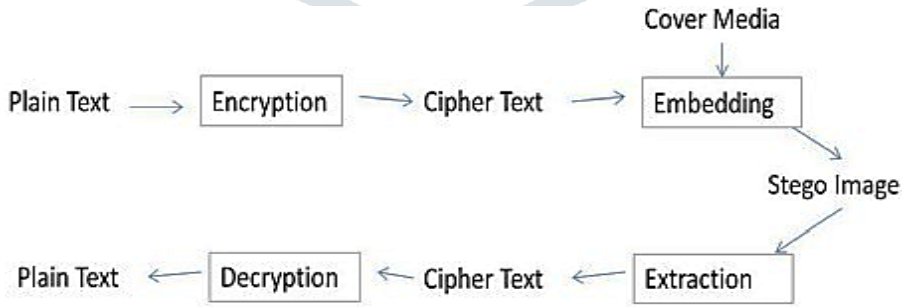

Figure 4 : Flow of Steganography + Cryptography

STEGANOGRAPHY + WATERMARKING

Watermarking is a technique that can be used to preserve the document's authenticity. Using a stego-key, this watermarked document can be inserted in a cover picture and delivered via the communication medium. The information can be decrypted using the reverse technique at the receiver end, and then its validity can be verified using watermarking. All four goals of data hiding will be met with this approach: security, capacity, robustness, and perceptibility.

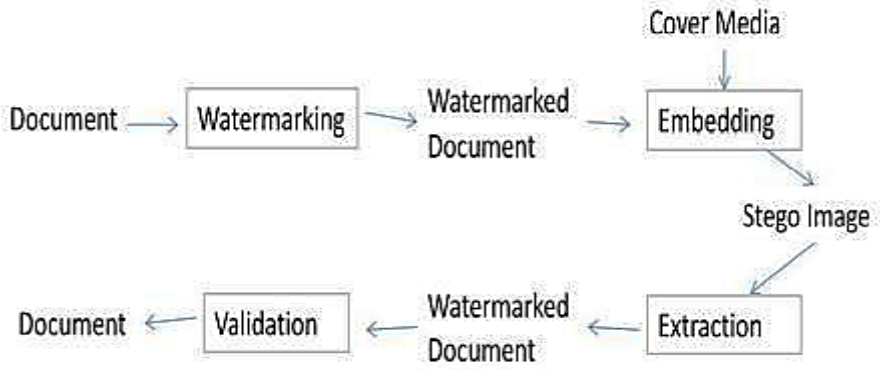

Figure 5 : Flow of Steganography + Watermarking

CONCLUSION

We attempted to provide a review of existing data masking techniques, as well as their benefits and drawbacks, in this study. This study also explains why data hiding is becoming more popular these days, as well as the objectives that any data hiding approach must meet. We've also attempted to explain how the basic aims of data concealing can be met by combining one or more data-hiding approaches.

REFERENCES

[1] Wawale, S. N., & Dasgupta, P. A. (2014). Review of Data Hiding Techniques. *International Journal for Advance Research in Engineering and Technology*, *2*(2), 32-37.

[2] Sarkar, T., & Sanyal, S. (2014). Steganalysis: detecting LSB steganographic techniques. *arXiv preprint arXiv:1405.5119*.

[3] Thampi, S. M. (2004). Assistant professor, department of computer science & engineering, lbs college of engineering, kasaragod, kerala-671542, s. *India—Information Hiding Techniques: A Tutorial Review*, *ISTE-STTP on Network Security & Cryptography, LBSCE*.

[4] https://en.wikipedia.org/wiki/Watermark

[5] https://en.wikipedia.org/wiki/Digital_watermarking

[6] Saranya, K., Mohanapriya, R., & Udhayan, J. (2014). A review on symmetric key encryption techniques in cryptography. *International Journal of Science, Engineering and Technology Research (IJSETR)*, *3*(3), 539-544.

[7] https://en.wikipedia.org/wiki/Steganography