Advancements in Steganography Techniques: A Comprehensive Review and Analysis

Abstract:

Steganography, derived from the Greek words "steganos" (meaning "covered") and "graphein" (meaning "writing"), is a fascinating technique that involves hiding sensitive information within seemingly inconspicuous carrier mediums. The primary objective of steganography is to ensure the secrecy and confidentiality of the hidden data while minimizing the chances of detection by unauthorized parties. This abstract aims to provide a comprehensive overview of steganography, including its fundamental principles, common methods, and applications across various domains.

At its corez, steganography exploits the imperceptible modifications that can be introduced into a carrier medium without raising suspicion. A carrier medium can refer to a diverse range of digital files, such as images, audio files, videos, or even text documents. By embedding secret information into these carriers, steganography allows for covert communication and secure data transmission.

Several techniques are employed in steganography to achieve effective concealment. One widely used method is the least significant bit (LSB) substitution, where the least significant bits of the carrier's pixel values are altered to encode the hidden information. This technique takes advantage of the human visual system's limited sensitivity to minor changes in pixel values. Spread spectrum modulation is another approach wherein the secret data is spread over a wide range of frequencies in an audio signal or other carrier types. Phase coding involves manipulating the phase information in the carrier to embed the concealed message.

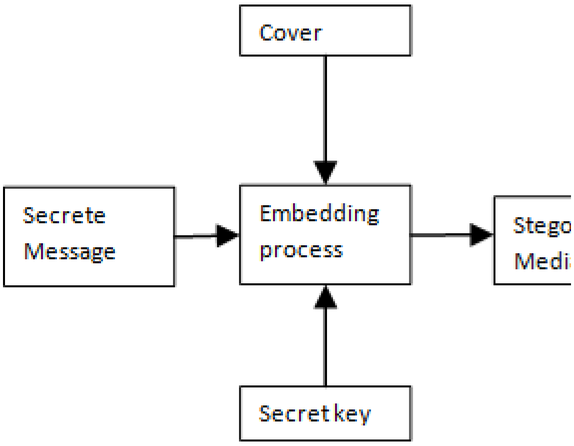
The applications of steganography are extensive and span diverse fields. In covert communication, steganographic techniques enable individuals or organizations to exchange confidential messages discreetly, ensuring privacy and security. These techniques have proven invaluable in scenarios where encryption alone might raise suspicion or draw unwanted attention. By embedding secret information within seemingly innocuous carrier files, steganography provides an additional layer of protection against prying eyes.

Keywords:

Steganography, Information hiding, Covert communication, Carrier mediums, Least significant bit substitution, Spread spectrum modulation, Phase coding, Digital watermarking, Encryption, Detection algorithms, Security, Privacy, Confidentiality, Data concealment, Multimedia steganography, Adaptive embedding, Machine learning-based detection, Multi-modal steganography, Robustness, Threats and countermeasures.

Introduction:

In today’s world, the communication is the basic necessity of every growing area. Everyone wants the secrecy and safety of their communicating data. In our daily life, we use many secure pathways like internet or telephone for transferring and sharing information, but it's not safe at a certain level. In order to share the information in a concealed manner two techniques could be used. These mechanisms are cryptography and steganography. In cryptography, the message is modified in an encrypted form with the help of encryption key which is known to sender and receiver only. The message cannot be accessed by anyone without using the encryption key. However, the transmission of encrypted message may easily arouse attacker’s suspicion, and the encrypted message may thus be intercepted, attacked or decrypted violently. In order to overcome the shortcomings of cryptographic techniques, steganography techniques have been developed. Steganography is the art and science of communicating in such a way that it hides the existence of the communication. The Greek word "steganography" means "concealed writing." The words "steganos" and "graphial" both mean "covered" and "writing," respectively. Steganography is therefore not only the technique of concealing data, but also the fact of transmitting secret data. The secret information is concealed using steganography in another file so that only the receiver is aware of its presence and message. In the past, information was safeguarded by hiding it on the scalps of slaves, the backs of writing tablets, the stomachs of rabbits, and wax. The majority of individuals nowadays, however, send data across the medium in the form of text, images, video, and audio.



Types of Steganography:

1. Text steganography: This technique involves concealing data inside text files. The secret data is concealed using this technique behind every nth letter of every text message word. There are numerous ways to hide data in text files.

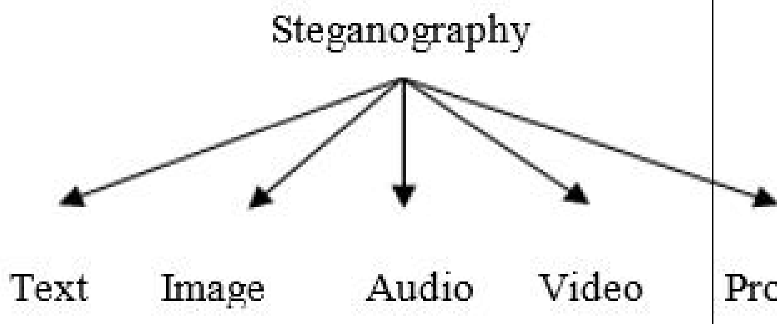
The first strategy is format-based, the second is random and statistical, and the third is linguistics-based.

2. Image Steganography: Image steganography is the practice of concealing data by using a cover object as an image. In image steganography, the data is concealed using pixel intensities.

3. Audio steganography entails the concealment of information in audio files. With this technique, the data in WAV, AU, and MP3 sound files is hidden. Different techniques exist for audio steganography. These techniques include phase coding and low bit encoding.

4. Video steganography: This method of data or file concealment in digital video format. In this instance, the data is concealed via video (a mixture of images). The data in each of the images in the video is typically hidden using discrete cosine transform (DCT), which typically modifies the values (e.g., 8.667 to 9) in a way that is invisible to the human eye. Video steganography uses the file types H.264, Mp4, MPEG, and AVI.

5. Network or Protocol Steganography: This method of information concealment uses a network protocol—such as TCP, UDP, ICMP, or IP—as a cover object. There are covert channels where steganography can be applied in the OSI layer network model.



JPEG File Compression:

The RGB color space is first converted into a YUV representation in order to compress an image into the JPEG format. The U and V components stand for color (or chrominance) while the Y component stands for brightness (or luminance) in this formulation. The human eye is known to be more sensitive to variations in pixel brightness than in pixel color [24]. When the color components (U and V) are split into their horizontal and vertical components, the JPEG is said to have the advantage of down sampling the color information, which reduces the file size by a factor.

The image is then changed. The discrete cosine transform (DCT) is utilized for JPEG images; using this mathematical procedure, the pixels can be changed by merely "spreading" the positions of the pixel values throughout the entire image or a portion of it. By grouping the pixels into (8 8) pixel blocks and then changing these blocks into 64-DCT coefficients, which are influenced by any variation of a single DCT coefficient, a signal is translated with DCT from the representation of an image into the frequency domain.

Image Generation Technique:

Many techniques have been proposed that encrypt messages so that they are unreadable or as secret as possible. Big Play Maker hides information by converting the secret text message into a larger and a slightly manipulated text format. The same principle can be employed in image creation, in which a message is converted to picture elements and then collected into a complete stego-image. This method cannot be broken by rotating or scaling the image, or by lossy compression. Parts of the message may be destroyed or lost because of cropping, but it is still possible to recover other parts of the message by encoding the message with error correcting information.

This method typically employs pseudo-random images because if an untrustworthy third party notices a collection of images moving through a network without any apparent reason (i.e., random images), he or she might assume that the images are carrying sensitive information and block their transmission [15].

File Embedding:

The header file architectures of various image file formats vary widely. Secret information can also be buried in a header structure or at the end of the file in addition to the data values, such as pixels, palettes, and DCT coefficients [47]. For instance, the header comment fields of JPEG images typically contain data that is concealed using Steganozorus and Secrets, both of which are undetectable. Data is added to the end of a JPEG image by camouflage, JpegX, PGE10, and PGE20.

The file header of image storage formats like TIFF, GIF, PNG, and WMF can be used to conceal any information. That random data in this situation might represent a covert message. To several picture storage formats, data can be added without changing the original image. Any tracking information at the file's end will be ignored when the image is prepared for display since the image user will decode the image size from the file header. This method enables you to add a message of any size to a cover image. However, the cover picture's message could be altered by simply saving the image again in the same file format [15].

Spread Spectrum:

Spread spectrum transmission in radio communications transmits messages below the noise level for any given frequency. When employed with steganography, spread spectrum either deals with the cover image as noise or tries to add pseudo-noise to the cover image. • Cover image as noise A system that treats the cover image as noise can add a single value to that cover image. This value must be transmitted below that noise level. This means that the channel capacity of the image changes significantly. Thus, while this value can be a real number, in practice, the difficulty in recovering a real number decreases the value to a single bit. To permit the transmission of more than one bit, the cover image has to be broken into sub images [15]. When these sub cover images are tiles, the technique is referred to as direct-sequence spread spectrum steganography. When the sub cover images consist of separate points distributed over the cover image, the technique is referred to as frequency-hopping spread-spectrum steganography. These techniques require searching the image for the carrier in order to then retrieve the data. These techniques are robust against gentle JPEG compression and can be made more robust through the pre-distortion of the carrier. In this case, after the carrier is created, and before the message is added, the carrier is compressed using JPEG compression and decompression such that it will be unaffected by later JPEG compression of the cover image [36]. The capacity can be traded directly for robustness, and it depends greatly on the image. • Pseudo-noise This technique shows that the hidden data is spread throughout the cover image and that is why it becomes difficult to detect [37]. Spread spectrum image steganography

(SSIS) described by Marvel et al., combined spread spectrum communication, error control coding, and image processing to hide information in images, is an example of this technique [38]. The general additive embedding scheme can be described as follows: Yi Xi Wi = + γ for i = ,2,1 ....,N (5) Where Xi is a sequence of the original data from the cover, Wi is a pseudo-random sequence generated from a pseudo-random number generator (PRNG) initialized by a secret stego key, γ is an embedding strength parameter (gain factor), and Yi is a sequence of possibly altered data. In SSIS, the process goes like this: the message is hidden in noise and then it is combined with the cover image to reach into a stego image. Since the power of the embedded signal is much lower than the power of the cover image, the embedded image becomes imperceptible not only to the human eye but also through computer analysis without access to the original image. The last few years witnessed the development of several steganography techniques one of which is spread spectrum steganography. In 1996, Smith and Comiskey described three schemes, namely direct sequence, frequency hopping, and chirp [36]. In image steganography, it is noticed that high frequencies usually aid the invisibility of the hidden information, but at the same time, they are not efficient as far as robustness is concerned. In contrast, low frequencies are better with respect to robustness, but are far too visible to be useful. Such conflicting points are reconciled by the spread spectrum technique via allowing the embedding of a low-energy signal in each one of the frequency bands, and as illustrated in [21].

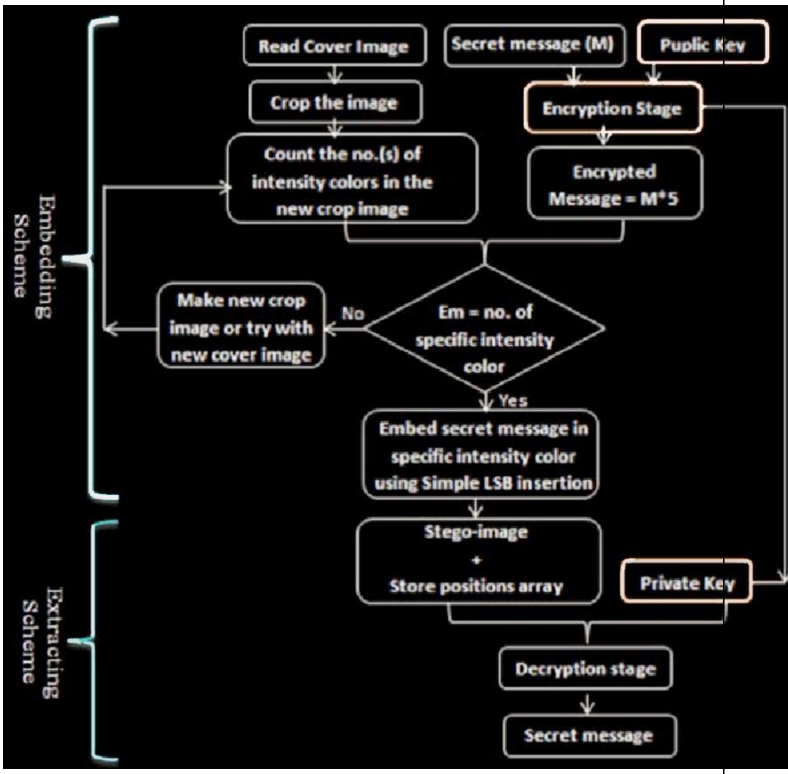
Adaptive Steganography:

A unique instance of the spatial and transform approaches is adaptive steganography. Additionally, statistics-aware embedding and masking is introduced. Prior to dealing with the frequency transformed coefficients of the image, the image's overall statistical properties are essentially utilised.

These statistics determine what adjustments are possible. This method, which relies on the cover picture and the selection of pixels in a block with a high standard deviation (STD), actually features a random adaptive selection of pixels. The latter is meant to stay away from smooth, uniformly colored surfaces. This method is notable for making use of photographs that have noise already present or that has been purposefully created, as well as images that exhibit complex color.

The "adaptive more surrounding pixels using" (A-MSPU) technique has been mentioned in [57] and it overcomes the imperceptibility issues with multiple base notational systems (MBNS). This method reexpresses the secret bits in several base notational systems while paying attention to the edge regions of a cover image. The suggested method scatters the secret bits using the same probability parameter and calculates each target pixel's capacity using the maximum number of neighboring pixels. The majority of steganographic methods employ the target pixel's three or four neighboring neighbors. All eight of the nearby neighbors can be used with the suggested method, increasing the imperceptibility value.

Methodology Diagram:


Read Cover Image
Secret message (M)
Puplic Key
Crop the image
Encryption Stage
Count the no.(s) of intensity colors in the new crop image
Encrypted Message = M*5
Embedding Scheme
Make new crop image or try with new cover image
No
Em = no. of specific intensity color
Yes
Embed secret message in specific intensity color using Simple LSB insertion
Stego-image + Store positions array
Private Key
Extracting Scheme
Decryption stage
Secret message

EVALUATION OF DIFFERENT TECHNIQUES:

- The first and most important criteria is undetectability (imperceptibility). This metric refers to the capacity to evade detection, i.e., where the human eye fails to see it. While not altering the image in a way that can be seen by the human eye, some techniques may nonetheless do so in a way that can be seen by statistical testing. Neither the naked eye nor statistical assaults should be able to discover really secure steganographic systems.
- Robustness: This second element gauges how well the steganographic approach can withstand attempts to decrypt the information being concealed. These efforts include data compression, image filtering, and image manipulation (such as cropping or rotation).
- They are not resistant to lossy compression and picture filters, and the problem of saving the image a second time completely ruins the hidden data [48]. Despite having a large payload, they are quickly found and defeated.
- Most often, attacks including cropping, rotating, and scaling, as well as those that target the watermarking approach, are successful against statistical methods. To make the statistical methods as reliable as the watermarking system, defenses could be taken into consideration. Depending on the cover image used, the payload capacity and invisibility may change.

Literature Review:

Steganography is a rapidly evolving field that has garnered significant attention from researchers due to its potential applications in secure communication, data protection, and information hiding. The literature review reveals that steganography techniques have been extensively studied and developed over the years, resulting in a diverse range of methods and algorithms for concealing information within carrier mediums.

One prominent area of research in steganography is the embedding process, where the hidden data is inserted into the carrier. Several techniques have been proposed, including the least significant bit (LSB) substitution, which involves modifying the least significant bits of pixel values in images or audio samples. This method provides a simple and effective way to hide information but may be susceptible to statistical analysis attacks. To counter this vulnerability, researchers have explored more advanced approaches such as spread spectrum modulation, which spreads the hidden data across multiple frequencies or time periods, making it harder to detect.

The literature review also highlights the importance of evaluating the security and robustness of steganographic methods. Researchers have proposed various metrics and evaluation criteria to assess the performance of different techniques. These metrics include capacity (the amount of data that can be hidden), imperceptibility (the degree to which the carrier is altered perceptually), and resistance to detection (the ability to withstand statistical analysis and other detection algorithms).

Another notable aspect of steganography research is the emergence of adaptive embedding strategies. These approaches aim to dynamically adjust the embedding process based on the specific characteristics of the carrier medium, optimizing the trade-off between imperceptibility and capacity. Adaptive embedding techniques leverage properties such as image complexity, audio characteristics, or text patterns to determine the optimal locations and strengths of data embedding.

Furthermore, advancements in machine learning have influenced steganography research. Researchers have explored the application of machine learning algorithms for both steganalysis (detection of hidden information) and steganography itself. Machine learning-based detection methods leverage pattern recognition and anomaly detection techniques to improve the accuracy and efficiency of detecting hidden data. On the other hand, machine learning models are also being utilized for enhancing steganographic algorithms by optimizing the embedding process based on learned patterns from large datasets.

The literature review reveals that steganography has found applications in various domains beyond covert communication, including digital watermarking and copyright protection. Digital watermarking employs steganography techniques to embed identifying information, such as a unique signature or copyright attribution, into media files. This allows for ownership verification and protection against unauthorized use or distribution.

In conclusion, the literature review demonstrates the extensive research and development efforts conducted in the field of steganography. It highlights the evolution of embedding techniques, evaluation metrics, adaptive strategies, machine learning applications, and diverse applications of steganography. The review emphasizes the importance of ongoing research to address challenges such as detection algorithms, security vulnerabilities, and the advancing complexities of digital media formats. By continuously exploring innovative approaches and adapting to emerging threats, researchers strive to enhance the security, robustness, and effectiveness of steganography methods in ensuring secure communication, data protection, and information hiding.

Conclusion:

We looked through numerous papers on steganography techniques for this study project. These studies are adequate and have a broad potential application. We discovered through reading these publications that the majority of the steganography work was completed in 2012 and 2013. These days, the most used steganography method is called LSB. In their studies, several researchers have also employed techniques including water marking, distortion, spatial, ISB, and MSB, which have produced a potent method of secure data transmission. The majority of the papers discussed here are sourced from publications like IEEE Explore, AICCSA, IJET, IJCSE, and IJCA. The initiator will benefit greatly from these papers when they begin their work in this area. Besides, file and spatial domain approaches are considered not to be robust against lossy compression and filtering. Transform domain techniques are considered more robust for lossy compression image formats, but this advantage is achieved at the expense of payload capacity. However, it is possible to defeat the transform domain techniques, but with some efforts. For most of steganography applications, JPEG file format can be used, especially for images that have to be communicated over an open systems environment like the Internet

References:

1. I.J. Cox, M.L. Bloom, J.A. Fridrich, and T. Kalkert. Digital watermarking and steganography. USA: Morgan Kaufman Publishers, 2008, pp. 1-591.

2. N. Provos and P. Honeyman. (2003, Jun.). “Hide and seek: An introduction to steganography.” IEEE Security and Privacy Journal. [On line], 1(3), pp. 32-44. Available: http://niels.xtdnet.nl/papers/practical.pdf [Jul., 2011].

3. S.B. Sadkhan. “Cryptography: Current status and future trends.” in Proc. IEEE Conference on Information & Communication Technologies, 2004, pp. 417-418.

4. N. F. Johnson, S. Katzenbeisser. “A Survey of steganographic techniques.” in Information Hiding Techniques for Steganography and Digital Watermarking, S. Katzenbeisser and F. Petitcolas, Ed. London: Artech House, 2000, pp. 43-78.

5. S. Areepongsa, N. Kaewkammerd, Y.F. Syed, and K.R. Rao. “Exploring on steganography for low bit rate Wavelet based coder in image retrieval system.” in Proc. of IEEE TENCON, 2000. pp. 250-255.

6. N.F. Johnson. (1995, Nov.). “Steganography. Technical report.” Available: http://www.jjtc.com/pub/tr_95_11_nfj/ [Sep., 2011].

7. P. Kruus, C. Scace, M. Heyman, and M. Mundy. (2003), “A survey of steganography techniques for image files.” Advanced Security Research Journal. [On line], 5(1), pp. 41-52. Available:

http://www.isso.sparta.com/documents/asrjv5.pdf#page=47 [Oct., 2011].

8. M. Kharazi, H.T. Sencar, and N. Memon. (2004, Apr.). "Image steganography: Concepts and practice." WSPC/Lecture Notes Series: 9in x 6in, [On line], pp. 1-49. Available:
http://iwearshorts.com/Mike/uploads/2011/06/10.1.1.62.8194.pdf [Aug. 2011].

9. D. Sellars. "An introduction to steganography. Internet:
http://www.cs.uct.ac.za/courses/CS400W/papers99/stego.html [Jul., 2011]+

10. M. Fortrini. "Steganography and digital watermarking: A global view." University of California, Davis. Available:
http://lia.deis.unibo.it/Courses/RetiDiCalcolatori/Progetti00/fortini/project.pdf . [June 2011].