Digital Innovations & Contemporary Research in SCIENCE, ENGINEERING & TECHNOLOGY

Vol. 10. No. 4, December 2022 – Special Issue on Cryptography & Security Mechanisms

Article Citation Format
Sackey, A.K. (2022): Steganography Based On Random Pixel Selection for Efficient Data Hiding. Journal of Digital Innovations & Contemporary Research in Science, Engineering & Technology. Vol. 10, No. 2. Pp 113-120
DOI: dx.doi.org/10.22624/AIMS/DIGITAL/V10N4P12

Steganography Based On Random Pixel Selection for Efficient Data Hiding

Sackey Abraham Kwaku
School of Technology
Ghana Institute of Management & Public Administration
GreenHills, Accra, Ghana
E-mail: abram.sackey@gmail.com
Phone: +233244153371

ABSTRACT

Advancement in technology has brought about the need for improvement in how data is stored and shared to prevent unauthorized access by third parties. Various methods including cryptography, watermarking, steganography and many others have been used in the quest to promote data safety. Currently, steganography has shown promise the most is being deployed by many. It simply involves hiding data in a media file but great success in data hiding has been chalked by this means. This paper explores the idea of steganography and its various types as well as how random pixel selection can be used to ensure efficient data hiding.

Keywords: Stenography, Random Pixel Selection, Efficient Data Hiding, Security

1. INTRODUCTION

Information safety has been one of the most important factors for information technology and communication since the rise of the internet (Morkel et al., 2005). Data hiding is the process of encapsulating information into media such as image and audio signals making it invisible to the human observer (Bender et al., 1995). The process is critical to ensure that even if the host signal is altered the data should continue to be buried within the host (Bender et al., 1995). It is important as it maintains the integrity of data and prevents infringement of copyright (Bender et al., 1995). Concealing data ensures that only authorized personnel have access to the data, preserving integrity by avoiding unauthorized substitutes. By minimizing interdependencies between software components, it also minimizes system complexity for improved robustness (Rakhi & Gawande, 2013).

Data hiding has been in existence since ancient times as relevant information was put on the back of wax, writing tables, stomach of rabbits or on the scalp of the slave (Rakhi & Gawande, 2013). It has however evolved as the emergence and advancements in technology has brought about information sharing through the internet (Wu et al., 2006). This means that "information thieves" have also resorted to the use of newer methods such as hacking for gaining unauthorized access to confidential data (Rakhi & Gawande, 2013; Wu et al., 2006). Steganography is one of the means of preventing such unlawful acts. Some other forms of data hiding systems includes digital watermarking, cryptography, fingerprinting and reversible data hiding (Morkel et al., 2005). Steganography is a data hiding technique similar to cryptography and watermarking. While watermarking ensures the message integrity and Cryptography scrambles a message, Steganography hides the message (Laskar & Hemachandran, 2013). The main difference between cryptography and steganography is that Steganography hides the existence of data while cryptography scrambles the data to change its meaning and quality so that it looks meaningless to others (Mahdi & Shafry, 2017).

Steganography is simply the art of rendering data imperceptible by hiding it in a media such as an image, a video, an audio or a text (Rakhi & Gawande, 2013). It is of Greek origin with the literal meaning of "covered writing". Unlike the forms of data hiding where the public have knowledge of the hidden information, imperceptibility is crucial to steganography (Morkel et al., 2005) It is able to hide data in plain such that there is little or no suspicion from anyone (Zhang et al., 2019).

Depending on the amount of data being hidden and the desired invariance to manipulation, different data concealing strategies are employed (Bender et al., 1995). Almost all digital file types can be used for steganography, however those with a high level of duplication are better ideal. The main file formats that are employed in this are texts, images, video/audio and protocol (Morkel et al., 2005). Of these texts are least used because they have a small amount of redundant data. The parts of an object that can be changed without the change being easily noticed are known as redundant bits (Morkel et al., 2005). Audio steganography makes use of the strategy that a faint sound is likely to be overlooked in the presence of a loud and a more audible one. In terms of redundant bits, it is comparable to image steganography, but because its sizes are larger than images, less of it is used (Morkel et al., 2005). Also, they are ubiquitous and easily accessible on the internet (Wu et al., 2006).

Steganography process consists of carrier, message, and password. Carrier is also known as cover object or cover-image, in which message is embedded. The message can be any type of data (plain text, cipher text, or image) that the sender wishes to remain confidential. Password has known as stego-key, which ensures that only recipient who has the stego-key will be able to extract the message from a stego-object. Finally, the cover-object with the secretly embedded message called the stego-object or stego-image. (MM Emam, AA Aly, FA Omara 2016)

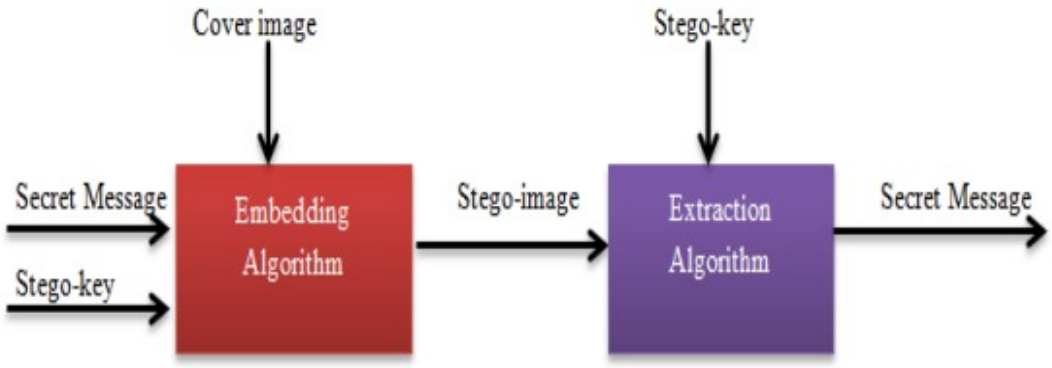

Fig. 1. The Basic Model Of Steganography

2. LITERATURE REVIEW

Overview Of Steganography

Steganography is the art and science of hiding a message inside another message without causing suspicion in others, so a message can only be detected by its intended recipient (Zaynalov et al., 2019). As stated earlier, the word steganography comes from the Greek language, "Stegos" meaning hidden or covered, and "Graphia" simply meaning covered writing (Laskar & Hemachandran, 2013). Steganography is a data hiding technique similar to cryptography and watermarking. When using cryptography, the presence of an encrypted message in itself attracts the attention of the attacker; in the case of steganography, the presence of hidden information remains imperceptible (Zaynalov et al., 2019). Steganography of digital image is a very useful tool for secure transmission of secret data, since it's very hard for human eyes to distinguish slightly modified stego image from a cover image (Mingwu et al., 2019). A common feature of steganographic methods and algorithms is that the hidden message is embedded in some innocuous, non-eye-catching object that is transported to the addressee openly (Zaynalov et al., 2019).

History and Wide Use of Steganography

The concept of hiding and concealing messages has existed for thousands of years, even though term steganography is only few years old (Swetha et al., 2015). Greeks were more successful in conveying secret messages (Amirtharajan & Rayappan, 2013). It is believed that steganography was first practiced during the Golden Age in Greece (Pooja & Kumar, 2010). The first steganographic technique was developed in ancient Greece around 440 B.C. (Siper et al., 2005). Herodotus, the Greek historian, reports how king Darius shaved the head of a prisoner and wrote a secret message on his scalp. After the hair grew back, the prisoner was sent to the king's son-in-law Aristogoras in Miletus and effectively delivering the message undetected by the enemy (Warkentin et al., 2008). In the same time period, another early form of steganography was employed. This method involved Demerstus, who wrote a message to the Spartans warning of eminent invasions from Xerxes. The message was carved on the wood of wax tablet, and then covered with a fresh layer of wax. This seemingly blank tablet was delivered with its hidden message successfully (Siper et al., 2005).

At the end of the Middle Ages, two authors produced seminal works on steganography. Johannes Trithemius (1462-1526) wrote the three volumes of Steganographia (ca. 1499) which superficially describe black magic, specifically using spirits to communicate over long distances. More than a century later, Gaspari Schotti picked up where Trithemius left off and published Steganographia (1665), which focuses on techniques with text, invisible inks, and incorporating hidden messages in music (Warkentin et al., 2008).

Other advancements included the advent of null ciphers, where unencrypted messages about ordinary events contain hidden messages during World War I (Warkentin et al., 2008). During World War II, invisible inks offered a common form of invisible writing. With the invisible ink, a seemingly innocent letter could contain a very different message written between the lines. Therefore, the text document can conceal a hidden message through using null ciphers (unencrypted message), which perfectly camouflage the real message in an ordinary letter (Por et al., 2014).

Invisible Ink involved common sources, this included milk, vinegar, fruit juice, and urine, for the hidden text. To decipher these hidden messages required light or heat (Siper et al., 2005). During World War II the Germans introduced microdots which were complete documents, pictures, and plans reduced in size to the size of a period and attached to common paperwork (Siper et al., 2005). The microdot technique was capable of hiding entire pages of text and even photographs, making them a powerful container of covert information (Swetha et al., 2015).

The use of steganography is now well within the reach of an average person with a computer and an Internet connection, and the most recent development is the potential use of steganography in Internet Telephony systems such as Skype (Warkentin et al., 2008). Present day steganography aims at only one thing “security in all means” and as internet has become necessity rather than luxury these days, the more information we access and share via the internet, the more risky the process is (Amirtharajan & Rayappan, 2013). Due to interception and improper manipulation by eavesdropper, data transmission in public communication system is not secure therefore an attractive solution for this problem is Steganography (Swetha et al., 2015).

There are many criteria for classifying steganography with one of them being classification based on the type of cover object (Swetha et al., 2015). The classification is as follows:

- Text steganography: The text steganography is a method of using written natural language to conceal a secret message (Swetha et al., 2015). In text steganography, different steganographic approaches are hiding by selection, HTML documents, line and word shifting, hiding using white space, semantic based hiding and abbreviation based hiding (Amirtharajan & Rayappan, 2013).

- Audio steganography: the process of hiding the information in a digitized audio and making it unnoticeable to the human ear is called audio steganography (Amirtharajan & Rayappan, 2013). Due to popularity of voice over IP (VOIP), audio has become a significant cover medium (Swetha et al., 2015). Audio steganographic domains may be broadly classified as Codec (bit stream hiding and modifying code book), Temporal (silence intervals, echo hiding and LSB) and Transform (wavelet, tone insertion, magnitude spectrum, cepstral domain techniques, etc) (Amirtharajan & Rayappan, 2013).

- Video steganography: Video Steganography is a technique to hide any kind of files in any extension or information into digital video format (Swetha et al., 2015). Since video is a combination of image and audio, their steganographic techniques are pertinent to video as well (Amirtharajan & Rayappan, 2013).

- Network/ Protocol Steganography: the phenomenon by which the secret information is rooted inside network protocols is called network/ protocol steganography (Amirtharajan & Rayappan, 2013). It refers to embedding information within network protocols such as TCP/IP, UDP, ICMP etc (Swetha et al., 2015).

- Image steganography: image steganography is the most accepted and widely used steganographic means in which information is transmitted from the sender to the receiver through innocuous interface (Amirtharajan & Rayappan, 2013). Images are used as the popular cover medium for steganography (Swetha et al., 2015). Image steganography techniques can be divided into Spatial Domain Methods, Transform Domain Technique, Distortion Techniques, Masking and Filtering (Swetha et al., 2015). There are many versions of spatial steganography (including Least significant bit (LSB), Pixel value differencing (PVD), Edges based data embedding (EBE), Quantization index modulation(QIM), Random pixel embedding (RPE), Pixel Mapping method, Multiple-Based Notational System, Difference Expansion Technique, Gray level modification (GLM), Labelling or connectivity method, Pixel intensity based method, Texture based method, Histogram shifting methods) all directly change some bits in the image pixel values in hiding data (Swetha et al., 2015).

3. RANDOM PIXEL SELECTION

Digital images often have a large amount of redundant data and the steganographic methods use this redundant data to hide the desired information. This is relatively easy because an image, being an array of pixels, typically contains an enormous amount of redundant information (Laskar & Hemachandran, 2013).

Least Significant Bit (LSB) steganography is where information is hidden in the spatial domain of an image and is the most broadly utilized steganographic strategy because of its effortlessness and clear methodology (Saha et al., 2020). It works by using the least significant bits of an image to hide the most significant bits of another (Laskar & Hemachandran, 2013). A simple example of LSB: Take this straightforward bit sequence as a piece of a carrier file:

10010101 00001101 11001001 10010110 00001111 11001011 10011111 00010000

Underlined are the Least Significant Bits in each byte group. The significance of these bits is so minor when compared to the whole, that altering these bits could produce close to the same result. 1001010**0** 0000110**1** 1100100**0** 1001011**0** 0000111**0** 1100101**1** 1001111**1** 0001000**1** (Siper et al., 2005)

Images consist of pixels with contributions from primary colours (red, green, and blue) adding to the total colour composition of the pixel and each pixel typically has three numbers associated with it, one each for red, green, and blue intensities, and these value often range from 0-255 (Laskar & Hemachandran, 2013). Depending on the depth of colour desired in the final image, each component is represented by a separate number of bits and by mixing the contribution of each component a large palette of colours can be represented (Laskar & Hemachandran, 2013).

The color of a pixel in an icon can be changed imperceptibly by minimally changing the digital color code (e.g. from 01011011 01010011 01011001 to 01011010 01010011 01011001) (Warkentin et al., 2008). When any specific colour is viewed closely, single digit modifications to the contribution level are imperceptible to the human eye. (i.e. a pixel with a value of (255, 255, 0) is indistinguishable from (254, 255, 0) (Laskar & Hemachandran, 2013).

Information bits can be embedded in image's LSB sequentially or ran domly distributed in image pixels (Amirtharajan & Rayappan, 2013). In sequential embedding approach, the LSB's of the image is replaced by the message bit sequentially or successively putting it at a disadvantage as the message is encoded in the image file sequentially (ordered), so we can find clusters of bits embedded, resulting in abrupt changes in the bits statistics, and this makes the detection easier (Laskar & Hemachandran, 2013). In random embedding of data, there are no such clusters because the embedded bits are scattered randomly in the image, so we cannot expect the easy detection process as compared to sequential embedding (Laskar & Hemachandran, 2013).

All steganography applications that use some kind of randomization technique in which the altered Least Significant Bits are spread out randomly across the carrier file creates the biggest obstacle for steganalysis (Siper et al., 2005). In the random embedding, the message bits are randomly scattered throughout the whole image using a random sequence to control the embedding sequence. The key used to generate pseudorandom numbers which is shared by both the sender and receiver, which will identify where, and in what order the hidden message is laid out (Laskar & Hemachandran, 2013)

4. CONCLUSION

The need for data hiding has increased with the rise of the internet which has become more of a necessity than luxury. Many people as well as organizations are looking for safe ways to send information back and forth while maintaining high level of privacy. This has equally given rise to steganography which is hiding information in plain sight and is more preferred to cryptography as it does not given an attacker the impression of something valuable being available. Image steganography is the most accepted and widely used steganographic means as it affords a larger amount of data to be hidden. This is because it contains an enormous amount of redundant information in an array of pixels. Least Significant Bit is the most popular technique in image steganography and involves embedding of information in the least significant bit of an image. Random pixel selection refers to randomly selecting bit of an image to embed information in. this technique decreases the changes of the information being noticed. This increases the level of security and efficiency of data hiding.

REFERENCES

1. Amirtharajan, R., & Rayappan, J. B. B. (2013). Steganography-Time to Time : A Review. *Research Journal of Informaton Technology*, *5*(2), 53–66. https://doi.org/10.3923/rjit.2013.53.66
2. Bender, W. R., Gruhl, D., & Morimoto, N. (1995). Techniques for data hiding. *Storage and Retrieval for Image and Video Databases III*, *2420*. https://doi.org/10.1117/12.205315
3. Laskar, S. A., & Hemachandran, K. (2013). STEGANOGRAPHY BASED ON RANDOM PIXEL SELECTION FOR EFFICIENT DATA HIDING. *Internation Journal of Computer Engineering & Technology (IJCET)*, *4*(2), 31–44.
4. Mahdi, M., & Shafry, M. (2017). Image Steganography Based on Odd/Even Pixels Distribution Scheme and Two Parameters Random Function. *Journal of Theoretical and Applied Information Technology*, *95*(22). www.jatit.org
5. Mingwu, Z., Zhang, S., & Harn, L. (2019). An efficient and adaptive data-hiding scheme based on secure random matrix. *PLoS ONE*, *14*(10), 1–14. https://doi.org/10.1371/journal.pone.0222892
6. Morkel, T., Eloff, J. H. P., & Olivier M S. (2005). An overview of image steganography. *Proceedings of the Fifth Annual Information Security South Africa Conference*.
7. Pooja, K., & Kumar, A. (2010). Steganography- A Data Hiding Technique. *International Journal of Computer Applications*, *9*(7).
8. Por, Y. L., Delina, B. M. Y., & Ang, T. F. (2014). WhiteSteg : A new scheme in information hiding using text steganography. *ReseachGate*.
9. Rakhi, & Gawande, S. (2013). A REVIEW ON STEGANOGRAPHY METHODS. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, *2*(10), 2278–2885. www.ijareeie.com
10. Saha, H., Yasmin, S., Chowdhury, R., Saha, G. C., & Masum, B. (2020). Random Pixel Selection Based Improved LSB Image Steganography Method Using 1 D Logistic Map and AES Encryption Algorithm. *International Journal of Innovative Science and Research Technology*, *5*(2).
11. Siper, A., Farley, R., & Lombardo, C. (2005). The Rise of Steganography. *Faculty Research Day*, 1–7.
12. Swetha, V., Prajith, V., & Kshema, V. (2015). Data Hiding Using Video Steganography -A Survey. *International Journal of Computer Science Engineering and Technology*, *5*(6), 206–213.
13. Warkentin, M., Bekkering, E., & Schmidt, M. B. (2008). Steganography : Forensic, Security, and Legal Issues. *Journal of Digital Forensics, Security and Law*, *3*(2).
14. Wu, N.-I., Wang, C.-M., & Hwang, M.-S. (2006). *Data Hiding: Current Status and Key Issues **.
15. Zaynalov, N. R., Kh, N. U., Muhamadiev, A. N., Bekmurodov, U. B., & Mavlonov, O. N. (2019). Features of using Invisible Signs in the Word Environment for Hiding Data. *Internation Journal of Innovative Technology and Exploring Engineering (IJITEE)*, *8*(9S3), 1377–1379. https://doi.org/10.35940/ijitee.I3295.0789S319
16. Zhang, M., Zhang, S., & Harn, L. (2019). An efficient and adaptive data-hiding scheme based on secure random matrix. *PLoS ONE*, *14*(10). https://doi.org/10.1371/journal.pone.0222892