Digital Innovations
& Contemporary Research in SCIENCE,
ENGINEERING & TECHNOLOGY
Vol. 10. No. 4, December 2022 – Special Issue on Cryptography & Security Mechanisms

Article Citation Format
Kwaku, T. (2022): Steganalysis Method for LSB Replacement Based
On Local Gradient of Image Histogram. Journal of Digital Innovations
& Contemporary Research in Science, Engineering & Technology.
Vol. 10, No. 2. Pp 1-8
DOI: dx.doi.org/10.22624/AIMS/DIGITAL/V10N4P1

Steganalysis Method for LSB Replacement Based On Local Gradient of Image Histogram

Kwaku, T.
School of Technology
Ghana Institute of Management & Public Administration
GreenHills, Accra, Ghana
E-mail: Kwaku.timothy@st.gimpa.edu.gh
Phone: +234244287747

ABSTRACT

Today's encryption and decryption processes rely heavily on the use of modern stenography techniques, which ultimately lead to steganalysis. Because of its enormous benefits, such as the protection of information or the transport of data, steganalysis is increasingly becoming an interesting research topic. This investigation is primarily concerned with steganalysis techniques, more specifically those that are utilized on photographic images. In this research, the many LSB steganalysis processes that are supported by previous research are investigated further, along with the outcomes of those operations. According to the findings of this research, in order for modern steganography and steganalysis techniques to be more efficient, they need to fundamentally incorporate artificial intelligence and machine learning methodologies.

Keywords: Steganalysis, LSB, Replacement, Local Gradient, Image Histogram

1. INTRODUCTION

Cryptography, watermarking, and steganography are the essential components of information concealment in the modern day; yet, each of these techniques seeks to accomplish a unique goal when it comes to their respective functions. The study of processing digital data by scrambling or encrypting in data bits with a key in such a manner that the data is unintelligible to an unauthorized person who does not possess the key to recover or decode it is known as cryptography. Cryptography is the study of how this may be done. When it comes to cryptography, it is abundantly evident that decrypting encrypted data that is either being sent or held in the memory takes an unacceptable amount of time and computer processing resources over the entirety of the data's useful life. Message data, however, can always be delivered in plain form without any restrictions, even if the customer who approved the distribution is permitted to do so. In addition, encryption unmistakably identifies a communication as one that contains information of interest, and thus makes the encrypted message vulnerable to being attacked (JinaChanu et al., 2012).

Steganalysis is largely related to the concept of embedded messages and process of determining whether a message under transmission has undergone some form of encryption. Steganalysis is primarily derived from the principle of steganography which essentially refers to the practice of passing on confidential information by means of seemingly harmless cover carriers in such a way that it is impossible to determine whether or not the cover carrier contains any hidden information.

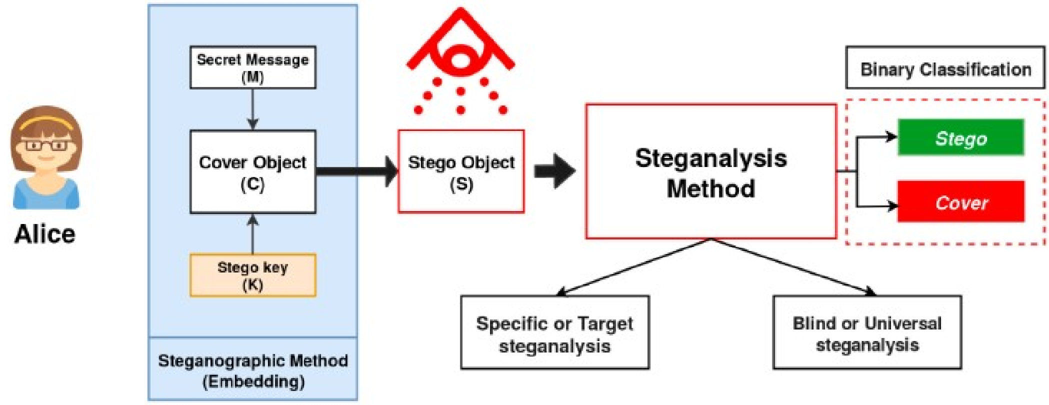

Fig 1: Typical Steganalysis Scenario
Source: https://www.linkedin.com/pulse/what-steganalysis-mayuri-bhamare/

These communications might be disguised as seemingly innocuous photos, music, video, or text files, or they could take the form of any other digitally represented code or transmission. The secret message might be in the form of plaintext, ciphertext, or anything else that is capable of being represented as a bit stream (Johnson & Jajodia, 1998). There are many different ways to conceal information in visual representations. Least Significant Bit (LSB) or noise insertions, alteration of picture and compression techniques, and altering of image attributes such as brightness are some examples of these approaches.

Steganography and steganalysis are complementary issues that typically arise in pairs and have drawn attention from all around the world. According to the adversary, the goal of steganalysis is to find any secret information that has been concealed in digital media. Its main need is the ability to reliably determine whether or not the test item contains sensitive information. To be more specific, it is even feasible to identify the type of steganographic method, calculate the message length, and extract sensitive information (Zhou et al., 2021).

Research Objectives

The underlying aim of this paper is to provide a comprehensive perspective grounded in literature, the steganalysis methods that may aid LSB replacements by means of local gradient of image histogram. By so doing this paper will:

1. Survey a list of steganalysis methods in modern day encryption from various literature sources
2. Identify and explore which of them (steganography and steganalysis) for LSB replacement based on local gradient of image histogram
3. Suggest recommendations for future works.

2. RELATED WORKS

According to Zhou et al., 3-D meshes are frequently employed to depict virtual surfaces and volumes. 3-D meshes have been popular in the last ten years for use in industrial, medical, and entertainment applications. They are extremely useful for 3-D mesh steganography and steganalysis. They conducted a thorough literature review on 3-D mesh steganography and steganalysis for their work. They suggest a new taxonomy of steganographic algorithms with four categories, as compared to a previous survey: 1) two-state domain, 2) LSB domain, 3) permutation domain, and 4) transform domain. They separate the steganalysis algorithms into two groups: universal steganalysis and customised steganalysis. The history of technological advancements and the state of technology now are introduced and discussed for each area. Finally, they discuss several attractive areas for future study as well as difficulties in enhancing 3-D mesh steganography and steganalysis performance (Zhou et al., 2021).

Sethi & Kapoor, 2016 projected a system that employs cryptographic method in addition to Steganography for appealing the defence of data thrashing and communication across networks. The file we wish to secure is first compressed to make it smaller, and then the data from that compression is converted into cypher text using the AES cryptographic technique, and finally the encrypted data is hidden in the image. In order to make it harder to discover secret information, genetic algorithms are employed for pixel selection in images where data must be hidden(Sethi & Kapoor, 2016).

The use of multimedia, including video, audio, and picture, has become quite common due to the quick growth of digital multimedia and online technologies. The problem of copyright security has recently drawn more and more attention as the number of applications rises. Steganography is a method for resolving this issue. Focus is placed on public key steganography in Shyla et al., 2016. Their research compares the performance of three alternative public key steganography techniques: the Edge Adaptive Scheme, the Multibit Assignment Scheme, and the Low Distortion Transform Scheme. The PSNR (Peak Signal-to-Noise Ratio), Embedding Capacity, and Embedding Time measurements are used to compare the performance of these approaches. According to the comparison research, the public key steganography system performs better than others while using Low Distortion Transform Shyla et al., 2016.

Yang et al. (2012) offer a novel weighted stego-image (WS) steganalysis approach to evaluate the ratio of messages concealed into each bit plane for concealing messages into multiple least significant bit (MLSB) planes. First, a novel WS with multiple weights is built, and it is demonstrated that the weight parameters are equal to the embedding ratios in MLSB planes when the squared Euclidean distance between the WS and the cover picture is low. A straightforward estimation equation is then established to estimate the embedding ratio in each bit plane based on this result and an estimated of the cover picture. The novel steganalysis approach beats the traditional structural steganalysis method on estimate accuracy when the embedding ratio in any bit plane is more than 0.4, according to experimental data. It also performs more steadily with changes in embedding ratios (Yang et al., 2012).

According to Shukla (2017), the development of cyber technology over the years has made data sharing quick and simple. Any type of data can be considered open data, including text, images, audio, and video. Data must be transformed into a digital format before it can be transmitted over the internet. People now have unrestricted access to the digital world thanks to the internet, but this might also pose serious security risks. Users must thus safeguard their data to prevent abuse. Cybersecurity ideas such as "data security in digital communication," "copyright protection of digitised properties," and "security of invisible communication via digital media" have been born as a result of this. The majority of steganalysis research has been conducted independently on targeted and blind steganography using different steganographic techniques. Although other steganalysis methods have been employed, the statistical steganalysis scheme has proven to be the most accurate of them. Since the steganographic process is known and the features that would obviously project the changes can be determined, the feature selection and extraction for focused steganalysis are made simpler. The differentiating characteristics for blind steganalysis must be recognised, chosen, and extracted depending on the picture format and image transformation (Shukla, 2017).

Ker et al. suggest a particular steganographic approach for LSB image matching (Ker, 2005). The Histogram Characteristic Function, or HCF, was initially developed (Harmsen & Pearlmana, 2003) for usage with colour photos. It is now being utilised with grayscale images. In order to use HCF, it was necessary to first calibrate the centre of mass (COM) with a down sampled picture and then compute an adjacency histogram rather than the traditional histogram. It was discovered that HCF-COM had a performance that was fairly good for colour photographs, but it turned out to have a performance that was rather terrible for grayscale images.

The first statistical steganalysis method was put out by (Westfeld & Pfitzmann, 2000). Pairs of Values (POVs) transferred during message embedding are identified using the approach. Pixel values, quantized DCT coefficients, or palette indices with different LSB values can all be used as POVs. According to Westfeld and Pfitzmann, each POV's two-pixel frequency values have a tendency to deviate greatly from the POV mean. These almost equal POVs in photographs and hence embedded information are discovered via the Chi-squared attack. When messages are randomly inserted, the Chi-squared approach identifies them less successfully than it does when they are implanted sequentially. (J. Fridrich et al., 2000) utilized a more extended version of the chi square assault to find messages that were randomly strewn throughout a picture.

The so-called Raw Quick Pair (RQP) approach was introduced by Fridrich et al. to identify LSB embedding in 24-bit colour pictures. RQP examines similar colour pairs produced by LSB embedding. Close colour pairings show that two colours only diverge at the LSB. The number of complementary hue combinations rises as a result of the process of embedding messages into visuals. Therefore, we may determine if a picture is stego or not by counting the number of close colour combinations. The authors demonstrated that it is feasible to attain a high level of detection reliability even for secret message capacities of 0.1–0.3 bits per pixel. This method's limitation to colour photos makes it a disadvantage. As a result of this, Jessica Fridrich et al., 2001 suggested a brand-new method known as RS steganalysis for detecting LSB embedding in colour and grayscale pictures. This method separates the image into groups and calculates the amount of noise in each group.

Each group is then classed as regular or singular depending on whether the pixel noise inside the group is raised or decreased by flipping the LSBs of a specified set of pixels within each group (by utilising a mask, i.e., the pattern of pixels to flip). Repeating the categorization results in a dual flipping type. Chi-square technique was shown to be less trustworthy than RS steganalysis.

Jessica Fridrich et al., 2003 presented the Pairs Analysis technique for 8-bit GIF pictures. The method makes use of colour slices, patterns made up of two colours, to determine how long the secret message is. They estimate the secret message length from the stego picture based on their measurement of the structure of the colour cuts using a variable R that resembles entropy but is really a quadratic function of the secret message length. This method works better than the Chi-square attack and produces more trustworthy results for BMP and palette pictures than RS steganalysis. However, Pairs Analysis performs somewhat worse for grayscale photos than RS steganalysis.

In their assault on a content-adaptive steganographic method (HUGO), (Jessica Fridrich et al., 2011) discovered traits that might detect payloads inserted using such techniques. They then used ensemble classifiers, which were created by combining the judgments of base learners who had been trained on random subspaces of the feature space. On the BOSSRank test set, the greatest performance was 80.3%, and the embedding rate was 0.4bpp. Hugo was also criticised by Gul et al. Prior to downsampling the picture, they first retrieved features by applying a function to the image that created estimates for the k variate probability density function (PDF). With an embedding rate of 0.4bpp, images from BOSSBase were utilised as the training set and those from BOSSRank as the test set. The detection accuracy was modestly increased by feature selection, on average by 0.3%. When 957 characteristics were chosen, an SVM was used as the classifier, and the best detection rate was obtained at 85%.

Mahdavi et al., 2008 introduced an innovative and accurate steganography analysis approach for LSB replacement steganography. The alterations that take place in the histogram of a picture following the embedding of data serve as the foundation for the proposed procedure. Every pair of nearby bins in a histogram are either inter-related or unrelated based on whether or not embedding a bit of data in the picture might impact both bins. This determines which of the two categories each pair of bins falls into. We show that an accurate measurement of the total quantity of embedded data may be obtained by comparing the overall behaviour of all inter-connected bins to the behaviour of bins that are not related to each other.

The correctness of the suggested technique has been demonstrated through both analytical analysis and simulation results. The recommended approach has been put into action, tested on more than two thousand samples, and compared to the RS Steganalysis method. When compared to the RS Steganalysis, the recommended approach had a mean error of 0.0025 and a variation of 0.0037, whereas the RS Steganalysis had a mean error of 0.0070 and a variance of 0.0182. They demonstrated, via the use of 4800 different examples, that the performance of the recommended technique is on par with that of the RS steganalysis for JPEG-filtered picture data. The novel method may be used to identify LSB embedding in either a random or sequential fashion depending on the context(Mahdavi et al., 2008).

3. METHODOLOGY

As was alluded to before, the primary focus of this article is a straightforward literature review that investigates a variety of steganographically relevant applications of various steganalysis approaches. In the methodology section of this study, highly standard journal databases were systematically explored in order to discover relevant material on the topic. IEEE Xplore, Google Scholar, the ACM digital library, Springer, and ScienceDirect were some of the databases that were featured in this list. The following are examples of some of the search words that were used: "Steganalysis definition," "Steganalysis procedures," "Steganography and Steganalysis applications," and other pertinent search phrases that meet the inclusion requirements of the methodology.

Inclusion and Exclusion Criteria

Literature and scenarios considered must necessarily be within the research scope were included. Thus, the papers considered must hit on either steganography, steganalysis, LSB replacement, Local Gradient of Image Histogram and other relevant topics relative to the underlying research goal. All papers that did not fall within the search scope were excluded from the work.

4. STEGANALYTIC TECHNIQUES

There are various steganalytic techniques that have been invented over the various years with the aim to decipher various encryption techniques. The science of steganalysis is the never-ending war against steganography. Active steganalysis employs specialized algorithms that recognise the presence of stego-image, whereas passive steganalysis changes image format, flips all LSBs, compresses JPEG, etc. in an effort to obscure the evidence of secret communication without bothering to discover the hidden message. Steganalysis may be divided into two groups: statistical and signature steganalysis. Both types of categories may be either particular or general. While universal steganalysis is a broad class steganalytic approach that may be used with any steganographic embedding algorithm, even an unidentified algorithm, specific steganalysis is developed for a specific steganographic embedding algorithm. JinaChanu et al., 2012, explained signature and statistical steganalysis as follows:

Signature Steganalysis

By inserting message bits that degrade or recur in patterns that serve as signatures for the presence of an embedded message, steganography modifies the characteristics of the media. A special signature is created by a steganographic technique like Hide & Seek that generates stego-images with pixel values that are divisible by 4, which mitigates the unsecure aspect for steganalytic instrument identification. The steganographic programme Jpegx works in a similar way by inserting a secret message before the JPEG file marker, followed by the hex number 5B 3B 31 53 00, which serves as a unique signature for detecting the secret message in the stego-image.

Statistical Steganalysis

Because mathematical procedures are more sensitive than visual perception, statistical steganography is a more potent kind of steganography than signature steganography. For the purpose of uncovering a hidden message from inside a stego-image that has been encrypted using LSB embedding, LSB matching, spread spectrum, BPCS, JPEG compression, or any other transform domain, specialised statistical steganalytic methods can be utilised.

The phrase "statistical steganalysis" refers to methods created by examining the embedding process and identifying certain statistics that are altered as a result of the embedding process. Because of this, achieving the highest level of steganalytic accuracy requires a thorough grasp of the embedding process. The steganographic technique is applied directly to the pixels of the picture in the spatial domain. Least Significant Bit Substitution, sometimes known as LSB, is one of the oldest methods. The LSB replacement and LSB matching LSB techniques were introduced as two distinct LSB methods (Karampidis et al., 2018).The Chi-square, RS, Gradient Energy-Flipping Rate Detection, and Histogram difference algorithms are among of the most powerful and widely used LSB detection methods.

Steganalysis on Images

Steganography and steganalysis have advanced significantly. Computer forensics is using steganalysis to pick and monitor documents suspected of uncontrolled occurrences to prevent internet data leaks. Each year, several new steganography methods are advocated, usually with enhanced steganalysis procedures for exposure. Due to the lack of information on the cover picture used to hide and recover data, cyber security steganalysis is a fascinating area. Targeted and blind steganalysis exist. Steganalysis for a given embedding procedure. The steganographic algorithm determines the differentiating traits needed for examination. Blind steganalysis is not embedding-dependent. It eliminates embedding technique performance dependence. Thus, this method works in many steganographic methods (Shukla, 2017).

LSB replacements

In LSB Replacement, the least significant bits of the cover image bytes are replaced with the secret data. In Least Significant Bit Substitution methods, there are two distinct embedding strategies, sequential and randomised. Sequential embedding indicates that the procedure begins at the first pixel of the cover picture and embeds the message data bits sequentially until the entire message is implanted. Randomized embedding scatters the locations of the values that will be updated to contain the embedded data bits.

5. RECOMMENDATIONS

Czaplewski, 2017, while focusing on blind steganalysis methodologies in the passive steganalysis scenario meant to discover the steganographic cover alteration procedures. The objective of this study is to explore the current state of the art in the subject of steganalysis and, more importantly, to identify the current trends that are present in this sector in order to establish some guidelines for the design of new steganalysis schemes. The consequences that are anticipated to occur are an investigation into the opportunities for the advancement of knowledge in the field of steganography and the formulation of guidelines for the conduct of further study (Czaplewski, 2017).

Deep learning algorithms and other AI technologies that can improve steganography methods for LSB substitution based on the local gradient of an image's histogram might be the subject of a significant amount of study. In further research, it may be possible to look into ways to enhance the performance of predicting the tiny embedding ratios and to study the possibility of applying the concept of WS to the steganography analysis of various adaptive steganography methods.

REFERENCES

1. Czaplewski, B. (2017). *Current trends in the field of steganalysis and guidelines for constructions of new steganalysis schemes Aktualne trendy w dziedzinie steganalizy oraz zalecenia dla konstrukcji*. *10*, 1121–1125. https://doi.org/10.15199/59.2017.10.3
2. Fridrich, J., Du, R., & Long, M. (2000). Staganalysis of LSB encoding in color images. *IEEE International Conference on Multi-Media and Expo*, *00*(III/WEDNESDAY), 1279–1282. https://doi.org/10.1109/icme.2000.871000
3. Fridrich, Jessica, Goljan, M., & Du, R. (2001). Reliable Detection of LSB Steganography in Grayscale and Color Images. *Proc. of the ACM Workshop on Multimedia Security*, 27–30.
4. Fridrich, Jessica, Goljan, M., & Soukal, D. (2003). *Higher-order statistical steganalysis of palette images*. *5020*, 178–190.
5. Fridrich, Jessica, Kodovský, J., Holub, V., & Goljan, M. (2011). Steganalysis of content-adaptive steganography in spatial domain. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6958 LNCS*, 102–117. https://doi.org/10.1007/978-3-642-24178-9_8
6. Harmsen, J. J., & Pearlmana, W. A. (2003). Steganalysis of Additive Noise Modelable. *Transform*, *2003*(April), 131–142.
7. JinaChanu, Y., Manglem Singh, K., & Tuithung, T. (2012). Image Steganography and Steganalysis: A Survey. *International Journal of Computer Applications*, *52*(2), 1–11. https://doi.org/10.5120/8171-1484
8. Johnson, N. F., & Jajodia, S. (1998). Steganalysis: The investigation of hidden information. *1998 IEEE Information Technology Conference: Information Environment for the Future, IT 1998*, *1998-September*, 113–116. https://doi.org/10.1109/IT.1998.713394
9. Karampidis, K., Kavallieratou, E., & Papadourakis, G. (2018). A review of image steganalysis techniques for digital forensics. *Journal of Information Security and Applications*, *40*, 217–235. https://doi.org/10.1016/j.jisa.2018.04.005
10. Ker, A. D. (2005). Steganalysis of LSB matching in grayscale images. *IEEE Signal Processing Letters*, *12*(6), 441–444. https://doi.org/10.1109/LSP.2005.847889
11. Mahdavi, M., Samavi, S., & Zaker, N. (2008). Steganalysis Method for LSB Replacement Based on Local Gradient of Image Histogram. *Iranian Journal of Electrical & Electronic Engineering*, *4*(3), 59–70.
12. Sethi, P., & Kapoor, V. (2016). A Proposed Novel Architecture for Information Hiding in Image Steganography by Using Genetic Algorithm and Cryptography. *Procedia Computer Science*, *87*, 61–66. https://doi.org/10.1016/j.procs.2016.05.127
13. Shukla, V. K. (2017). *AN OVERVIEW OF FEATURE BASED STEGANALYSIS*. *14*(2), 224–229.
14. Shyla, S. I., Adaptive, A. E., & Steganography, I. (2016). *Empirical Evaluation Of Image Steganography*. *3*(11), 74–76.
15. Westfeld, A., & Pfitzmann, A. (2000). *Attacks on Steganographic Systems*. 61–76. https://doi.org/10.1007/10719724_5
16. Yang, C., Liu, F., Lian, S., Luo, X., & Wang, D. (2012). Weighted stego-image steganalysis of messages hidden into each bit plane. *Computer Journal*, *55*(6), 717–727. https://doi.org/10.1093/comjnl/bxr112
17. Zhou, H., Zhang, W., Chen, K., Li, W., & Yu, N. (2021). Three-Dimensional Mesh Steganography and Steganalysis: A Review. *IEEE Transactions on Visualization and Computer Graphics*, *2626*(c), 1–20. https://doi.org/10.1109/TVCG.2021.3075136