

Application of Digital Steganography for Information Encryption

Akinrinlola Ibitoye Akinfolo, Amusu Mary

1.0 INTRODUCTION

Due to the recent advancement in technology across the globe, a lot of people now prefer to transfer data from one end to another end using the internet as the primary medium of transfer. There are now several ways which data transmission can take place on the internet, some of which includes: emails, social media, cloud sharing (e.g, Onedrive, Google drive) etc. Internet has made data transfer from one end to another an easy, rapid and precise task. Nonetheless, security threats have also emerged as one of the greatest problems associated with data transmission across the internet. For example, a confidential or sensitive document being transmitted from one person to the other can be stolen or destroyed by an attacker using several malicious strategies. So, it becomes very vital to give data security utmost attention because of the various inherent issues associated with the current emerging technologies used in the transmission of data.

Data security in its simplest form is an act of protecting data or information from unauthorized users or attackers. It also entails providing a preventive means against the modification of data by wrongdoers. This field of study has seen massive increase in its research lately due to the rate at which data transfer over the internet has grown exponentially. Several security features have been proposed by different researchers lately, some of which includes: Steganography, Cryptography, Digital watermarking, and so on. All these are to increase security of data transfer across the internet.

Cryptography in its own context is a methodology used to hide information by the means of encryption and transmitting it to its actual receiver using a secret key to encrypt it. Upon receiving the data, the receiver then uses the secret key to decrypt the data so that the real content can be revealed. Steganography is a methodology that further provides data security by hiding the cipher texts inside another file, hence making it seem no data is hidden at all.

Johnson et al., (2001), defined “Steganography as an art of concealing and transferring data through an innocent carrier to keep out of sight the presence of data”. The visibility of the hidden data is decreased by applying some hiding methodology such as: LSB, F4, Jsteg and so on. The act of uncovering or revealing the information concealed in these algorithms or methodology is called Steganalysis. The text which was encoded by

the technique is referred to as “Cipher text” and the technique itself is known as “encryption” and this encryption can be undone with an authorized access using decryption technique, in which encoded data is now transformed into a format that is readable (Kahate, 2008).

“Steganography” and “Cryptography” are associated paradigms. The concealed or embedded files represent the carriers which are used to transmit the messages to their intended destination without security scare or concerns. Steganography was initially implemented on image file format but there are now growing implementations which will extends to other media file formats. Nonetheless, images are mostly preferred for this technique. At the moment, there are many algorithms presently in use for developing steganography software (Krenn, 2004).

Digital watermarking is a tool which can help tackles the issue of copyright on digital platforms. It is majorly reliant on Steganographic method. It acts a very good solution to copyright problems by embedding a digital symbol like a watermark which is not manually editable. There is an important thing to keep in mind when using steganography, it is to ensure that the originality of the image after the data is embedded is not compromised.

Data is becoming more and more valuable daily. Hence, for salvaging diverse types of private data are growing rapidly. Improper handling or transfer of personal information can lead to great harm to individuals and to businesses. Like most other information security problems, the effect of not hiding sensitive information on its targeted victim includes:

- i) Loss of confidential information which can in turn be used to harm the owner in different ways.
- ii) Alteration of information for the benefit of the attacker.
- iii) Destruction of vital documents which can lead to loss of money for individuals or companies.

2.0 LITERATURE REVIEW

2.1. Review of Related Works

2.1.1 Information Communication Technologies (ICT)

According to Ratheeswari, K. (2018), Information communication technologies (ICT) at present are influencing every aspect of human life. They are playing salient roles in work places, business, education, and entertainment. Moreover, many people recognize ICTs as catalysts for change; change in working conditions, handling and exchanging information, teaching methods, learning approaches, scientific research, and in accessing information communication technologies. In this digital era, ICT use in the classroom is important for giving students opportunities to learn and apply the required 21st century skills. ICT improves teaching and learning and its importance for teachers in performing their role of creators of pedagogical environments. ICT helps of a teacher to present his teaching attractively and able to learn for the learners at any level of educational programmes. Today in India teaching training programmes making useful and attractive by the term of ICT. Information and Communication Technologies (ICTs) exemplified by the internet and interactive multimedia are obviously an important focus for future education and need to be effectively integrated into formal teaching and learning – especially in a teacher education institution.

2.1.2 Information Security

Moody, G. et al (2018), examined issues related to information systems security (ISS) behavioral research has produced different models to explain security policy compliance. This paper (1) reviews 11 theories that have served the majority of previous information security behavior models, (2) empirically compares these theories (Study 1), (3) proposes a unified model, called the unified model of information security policy compliance (UMISPC), which integrates elements across these extant theories, and (4) empirically tests the UMISPC in a new study (Study 2), which provided preliminary empirical support for the model. The 11 theories reviewed are (1) the theory of reasoned action, (2) neutralization techniques, (3) the health belief model, (4) the theory of planned behavior, (5) the theory of interpersonal behavior, (6) the protection motivation theory, (7) the extended protection motivation theory, (8) deterrence theory and rational choice theory, (9) the theory of self-regulation, (10) the extended parallel processing model, and (11) the control balance theory. The UMISPC is an initial step toward empirically examining the extent to which the existing models have similar and different constructs.

2.1.3 Image Encryption Algorithm

According to Shawn, D. (2017), to ensure security, image encryption algorithms generally include two stages: permutation and diffusion. The traditional image permutation algorithms which include the Sort-based permutation algorithm,

Arnold-based permutation algorithm, Baker-based permutation algorithm and the Cyclic Shift permutation algorithm, etc. However, these algorithms have the disadvantages of either high time complexity or poor permutation performance. Therefore, in combination with cyclic shift and sorting, his research proposed a permutation algorithm that cannot only guarantee good permutation performance but also guarantee low time and space complexity. Most importantly, the research proposes a parallel diffusion method. This method ensures the parallelism of diffusion to the utmost extent and achieves a qualitative improvement in efficiency over traditional streaming diffusion methods. Finally, combined with the proposed permutation and diffusion, the paper proposes a computational model for parallel image encryption algorithms.

Li, C. et al (2018), presented a paper that performed a thorough security analysis of a chaotic image encryption algorithm based on autoblocking and electrocardiography from the view point of modern cryptography. The algorithm used electrocardiography (ECG) signals to generate the initial key for a chaotic system and applies an autoblocking method to divide a plain image into blocks of certain sizes suitable for subsequent encryption. The designers claimed that the proposed algorithm is “strong and flexible enough for practical applications”. This became vulnerable to the known plaintext attack: based on one pair of a known plain-image and its corresponding cipher-image, an adversary is able to derive a mask image, which can be used as an equivalent secret key to successfully decrypt other cipher images encrypted under the same key with a non-negligible probability of $1/256$. Using this as a typical counterexample, some security defects existing in many image encryption algorithms were summarised.

2.1.4 **Steganography**

Steganography is a word which finds its origin in the Greek language and it simply means “covered writing”. Steganography is the method used to hide the existence of information by embedding one information/data into another such that the existence of the hidden data is not visible to the human sight. There are a number of techniques used to perform steganography depending on the carrier/cover data used.

Steganography and cryptography are related because both of them are used to implement security in files to be sent from one location to the other. Almost the same approach of using secret key used in cryptology is also employed in steganography. In steganography the data is hidden while in cryptography the data is transformed into an unreadable format.

Some existing types of steganography are:

- i) Pure steganography
- ii) Public key steganography
- iii) Secret key steganography

Pure Steganography: In pure Steganography, the files are embedded into a cover image without using any secret key and deciphering the embedded image as well and this will not require any secret key. The success of this method relies solely on the ignorance of a user about the existence of a hidden information.

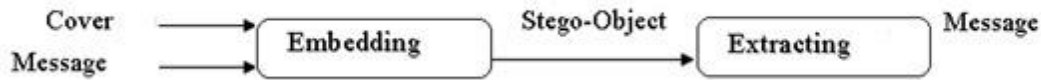


Figure 2.1: Pure Steganography Process (Zaidoon, 2010).

Pure Steganography is not so secure because the attackers can easily decrypt it if he is aware of how this algorithm works. The only advantage of it is that it does not require any key to be shared between the sender and the receiver.

Secret key Steganography: This uses the same approach as the pure steganography but includes the use of secret key for encryption and the same key will be required during decryption of such information.



Figure 2.2: Secret key Steganography (Zaidoon, 2010).

Secret key steganography comes with an improved security over the pure steganography algorithm. This method is very secure, but the only risk is when the attacker finds a way to extract the secret key from either the sender or the receiver.

Public key steganography: The public key steganography is similar to the secret key steganography, the only difference or improvement is that different keys are required for the encryption (private key) and decryption (public key) and it is usually stored in a public database for retrieval.

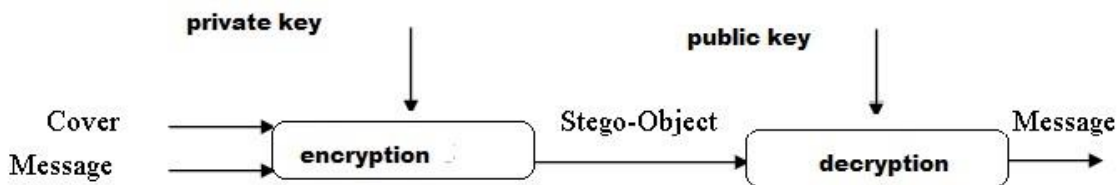


Figure 2.3: Public key Steganography (Zaidoon, 2010).

IJSER

The regularly used methodologies for implanting data into images are

- i) LSB (Least Significant Bit) Algorithm and
- ii) JSteg Algorithm

Transmitting confidential images between two channels suffer from hacking. Therefore, protecting confidentiality is a very essential issue. Recently, several methods have developed to protect important information. The main idea is based on embedding important information in multimedia carrier such as: text, image, audio, and video. The developed methods may be classified as steganography and watermarking.

Steganography aims to embed huge amount of secret data in a multimedia carrier while watermarking aims to hid small amount of secret data in multimedia carrier. According to Alaa, F. et al (2016), in their research first presented a literature survey of information hiding, then classified the proposed methods, and finally introduces a comparative study between the different methods.

Bandyopadhyay S. K., & Maitra, I. K. (2010), in their research used an Alternative Approach to Steganography using Reference Image. The research created a practical steganographic implementation for 4-bit images. The proposed technique converts 4 bit image into 4 shaded Gray Scale image. This image will act as reference image to hide the text. Using this grey scale reference image any text can be hidden. Single character of a text can be represented by 8-bit. The 8-bit character can be split into 4X2 bit information. If the reference image and the data file are transmitted through network separately, we can achieve the effect of Steganography. Here, the image is not at all distorted because said image is only used for referencing. Any huge amount of text material can be hidden using a very small image. Decipher the text is not possible intercepting the image or data file separately. So, it is more secure.

2.1.5 Cryptography

The word cryptography springs from two Greek words which mean “secret writing”. Cryptography is that the process whereby an original text is scrambled through the replacement and rearrangement of the initial text, transforming it into an unreadable format for others. Cryptography is an efficient way to protect the knowledge that is transmitting through the network communication paths (Bishop, 2005).

Cryptology is the science that deals about cryptography and cryptanalysis. Cryptography is the approach of sending the messages secretly and securely to the destination. Cryptanalysis is the method of obtaining the embedded messages into original texts (Whitman, 2007).

Cryptography is the act of sending data from one source to another destination by altering its appearance by using a secret key. A cryptosystem makes use of a plain visible text as

its input to generate a cipher or transformed text by using an encoding algorithm using a secret code as input.

The basic elements of cryptosystems are:

- i) Plain text (input)
- ii) Encryption algorithm
- iii) Secret key
- iv) Cipher text
- v) Decryption algorithm

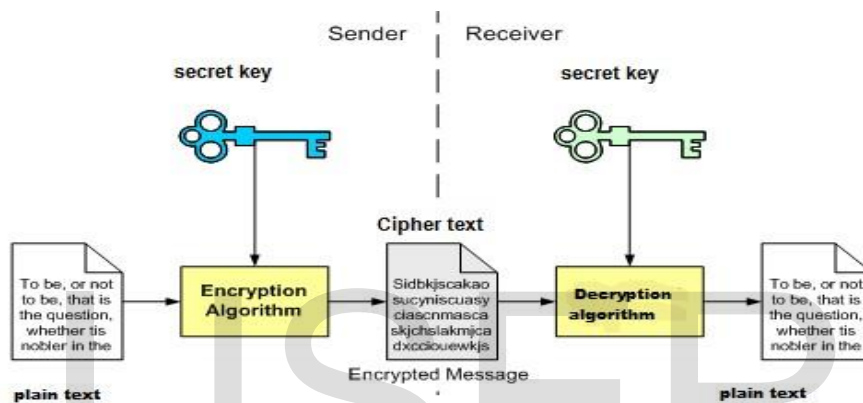


Figure 2.4: General model of Cryptographic System

Cryptographic Algorithms: This refers to various classes of encryption algorithm. Depending on the standard used for the encryption the algorithms are of two types namely:

- i) Symmetric encryption algorithm
- ii) Asymmetric encryption algorithm.

Symmetric Encryption

Symmetric encryption uses a single key encryption method. It is also called conventional encryption. The symmetric encryption algorithm uses same code for encrypting data and decrypting the data as well. What determines the level of security in this type of algorithm is the length of the key.

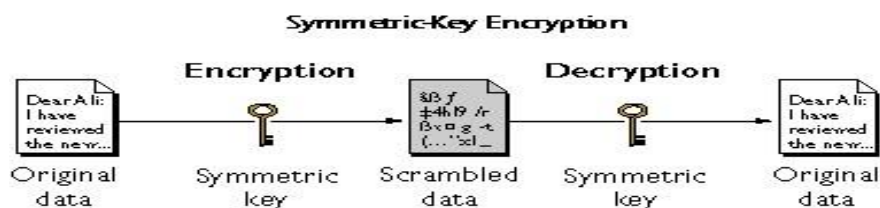


Figure 2.5: Symmetric encryption

Asymmetric Encryption

Asymmetric encryption or Public key encryption performs the same work as Symmetric encryption, but the major dissimilarity is their use of keys. Using asymmetric encryption, different keys are used in the decryption and encryption.

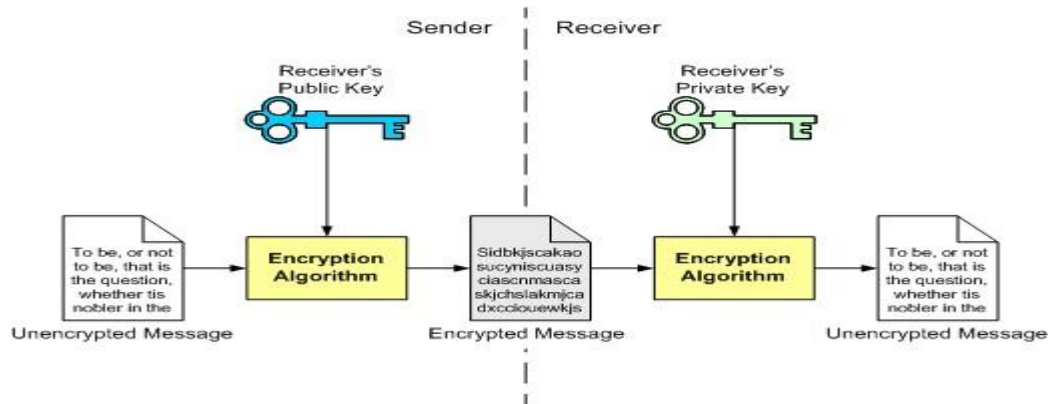


Figure 2.6: Asymmetric Encryption

In Asymmetric encryption, the only data which was encrypted with a public key is the only ones that can be decrypted with the same. And only a corresponding matching key with the one used for the encryption can be used for the decryption as well.

The major identified issue with Asymmetric algorithm is “cipher keys”. If a sender wants to transfer data to a receiver, they will both need up to four different keys. It will become more complicated because each of those keys must be the same at each of the various levels. RSA algorithm is sometimes regarded as the most vital public key encryption algorithm

2.3.1 Least Significant Bit (LSB) Algorithm

In Least Significant Bit (LSB) substitution, the steganography is performed by adjusting the least significant bit pixels on the cover media usually an image. It is a methodology used for implanting messages into an image. The LSB insertion differs given the number of bits in the cover/carrier image. LSB has been proven to be a highly effective algorithm for hiding messages inside an image provided the cover are large enough to contain the intended message to be hidden. There are many approaches available for hiding the data within an image: one of the simple least significant bit submission approaches is “Optimum Pixel Adjustment Procedure”.

The simple algorithm for OPA explains the procedure of hiding the sample text in an image.

Step1: A few least significant bits (LSB) are substituted with in data to be hidden.

Step2: The pixels are arranged in a manner of placing the hidden bits before the pixel of each cover image to minimize the errors.

Step3: Let n LSBs be substituted in each pixel.

Step4: Let d = decimal value of the pixel after the substitution.

d_1 = decimal value of last n bits of the pixel.

d_2 = decimal value of n bits hidden in that pixel.

Step5: If $(d_1 \sim d_2) \leq (2^n)/2$ then no adjustment is made in that pixel.

Else

Step6: If $(d_1 < d_2)$

$d = d - 2^n$.

If $(d_1 > d_2)$ $d = d + 2^n$.

This d is converted to binary and written back to pixel (Amirtharajan et al., 2010).

This method of substitution is simple and easy to retrieve the data and the image quality better so that it provides good security.

IJSER

3.0 RESEARCH METHODOLOGY

This has to do with the specification of procedures for collecting and analyzing data necessary to define or solve the problem for which the research is embarked upon. The scope of this research covers the concept of steganography and how it can be improved for securing data transmission from one location to the other.

The methods employed in gathering of data for this research work is mainly through the use of the internet to get a good knowledge of what has been done by other scholars in the past.

3.1.1 Primary Source

This involves oral interviews conducted with various lecturers and past students of the institutions who may have an idea as to what steganography does and how it works to provide security of data.

3.1.2 Secondary Source

In the course of this research the secondary sources that were used include:

- a. Dictionaries: Which were used to get the meaning of some new words and in the composition of the research.
- b. Textbooks: Some computer science related textbooks were also consulted for us to be able to understand some of the concepts of data security amongst other things.
- c. Journals: Various Journals were used consulted during the course of the research
- d. Other materials used include electronic books and internet downloads for collection of data and to aid comprehension of the research work.

For a steganographic algorithm to be successful it must meet the objectives listed below:

- i. **Invisibility:** A steganographic methodology must not be visible to the human eye, because its actual aim is to take off unwanted attention of hackers from the hidden information being transmitted. If an attacker's eye suspects that an image contains a hidden information, then the objective of the steganography is not met. This may lead to such data being compromised.
- ii. **Robustness against image manipulation:** A Steganography procedure is more desirable if it is effective against malicious or unanticipated changes to the image.
- iii. **Independent of file format:** The most common type of file being embedded in a steganographic technique is text file, but an effective steganography technique should be able to hide other file formats as well.

An important criterion of a good steganographic system is that the quality of the image being used (cover image / stego-image) for steganography purposes must be as close as possible to the original image, as not to raise suspicion or attract any unwanted attention to the stego image.

In the past few years, several steganography techniques that embed information into plain files such as image, audio, video etc. have been developed and employed. However, in the current image-based steganography technology, there are some problems identified with the system. Some of which are:

- i. Most of them focuses on embedding one file type (e.g. text or image) inside a cover file.
- ii. Distortion produced in a stego-image after inserting information to be hidden. Due to this factor, the stego-image produced can raise suspicion and can be clearly distinguished by an attacker. Thus, researches on alleviating the distortion problem in the stego-image need to be carried out.
- iii. Implementation of a single steganography technique that can hide all format of data into image is quite scarce to come by.
- iv. Some of the existing implementations are CLI (Command Line Interface) base which makes them unfriendly to users.
- v. The speed of performing the steganography as well is not fast in most of the existing system.

IJSER

4.0 RESULTS AND DISCUSSION

This section presents the screenshots and results of the implementation of the research.

4.1 Overview of the Proposed System

The proposed digital Steganography system will require any type of image file (such as jpeg, bmp, png etc.) and the information or message that is/are to be hidden. The image file is what is referred to as the Cover Image or Agent File and it is required mainly to act as the carrier of the actual data to be hidden. The information to be hidden can be in mainly any format (Such as image, texts files, audio or simply a message to that requires secret transmission). It will have three (3) modules which are:

- a. **Hiding a file:** This module will be used for embedding a file of several formats into a cover image. This module requires a cover image and another file to be hidden as the input and outputs the stego-image as an output.
- b. **Hiding a message:** This module is similar to the one above, but the only difference is that the user will be able to type and hide a message instead of a file. This module takes in the cover image and message as input and outputs a single file which is called the stego-image.
- c. **Revealing a file/message:** This module will be used to reveal or unhide a message that has once been hidden by the first or second module. This module takes the stego-image as an input and outputs the same file and the hidden message or file.

The Steganography technology will be written and implemented using the JAVA Programming Language and the Algorithm will be achieved using the LSB (Least Significant Bit) Algorithm. The algorithm used for hiding and unveiling this application provides several layers in lieu of using only LSB layer of image. Writing data starts from last layer (last or LSB layer); because significant of this layer is least and every upper layer has doubled significant from its down layer. So, in every step we go through, the upper layer image quality decreases and image retouching transpires.

i. Default Page

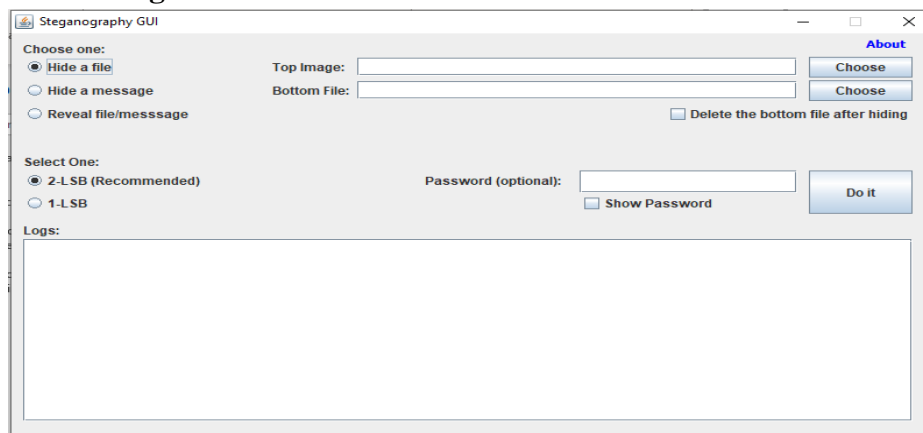


Figure 4.1: Default page

The above shows the main page where a user is required to select whether to hide a file, hide a message or reveal file/message. The user also has additional security on whether to include a password in the cipher process or not.

The 2-LSB and 1 LSB option has different output size result. Meaning that after hiding the file in a picture, the size of 1 LSB is usually smaller than 2 LSB.

ii. Top Image

The top image is also known as the cover image and it is required that a user selects an image file (such as .png, jpeg, etc) which will be used as the carrier of the actual information being protected. It is advised that the resolution of the image to be used as the top image be high to accommodate more data.

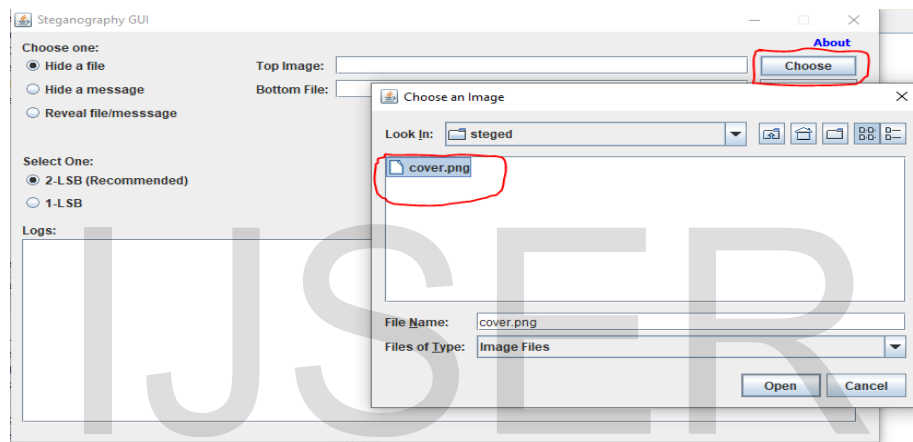


Figure 4.2: Top image

The user is expected to click Open after selecting a file to be used as the top image.

iii. Bottom File

The bottom file can be any kind of file at all which the user is intending to secure. In this example, we are hiding a pdf document inside a cover.png file.

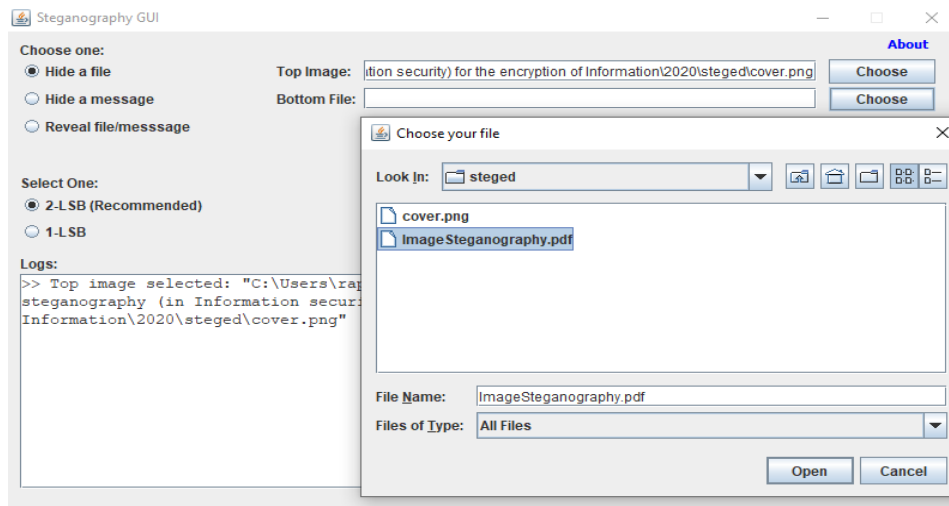


Figure 4.3: Bottom file

iv. **Delete Bottom File (Optional).**

The user can choose to delete the original file after it has been hidden in the top image. It is worth stating that files deletion in this case can be difficult to recover if the user forgets his/her stego key (a password used to hide the file).

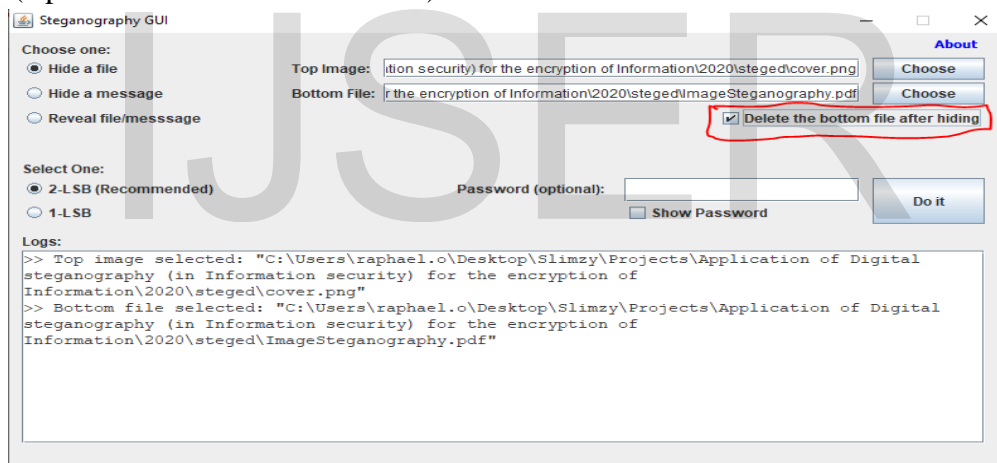


Figure 4.4: Delete Bottom File (Optional)

v. **Set a password:**

Setting a password is also optional as the file cipher will still take place even if no password is set.

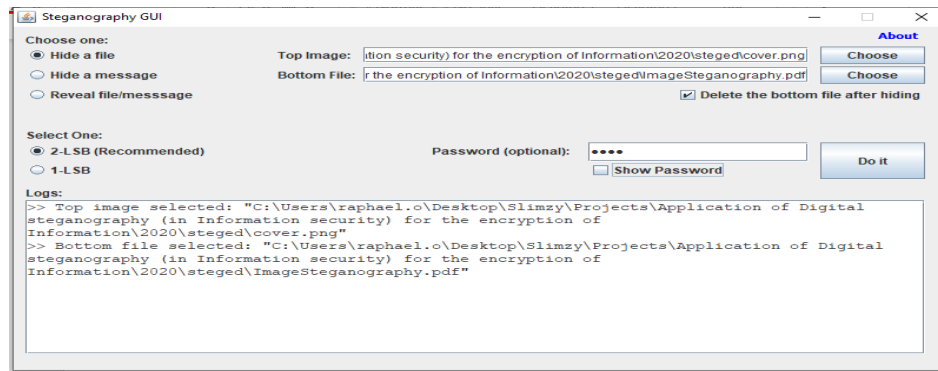


Figure 4.5: Set a password

Once the user completes the required steps above. He/She can click on “DO it”.

vi. **Save File**

The user is then required to specify the location where the stego image will be saved.

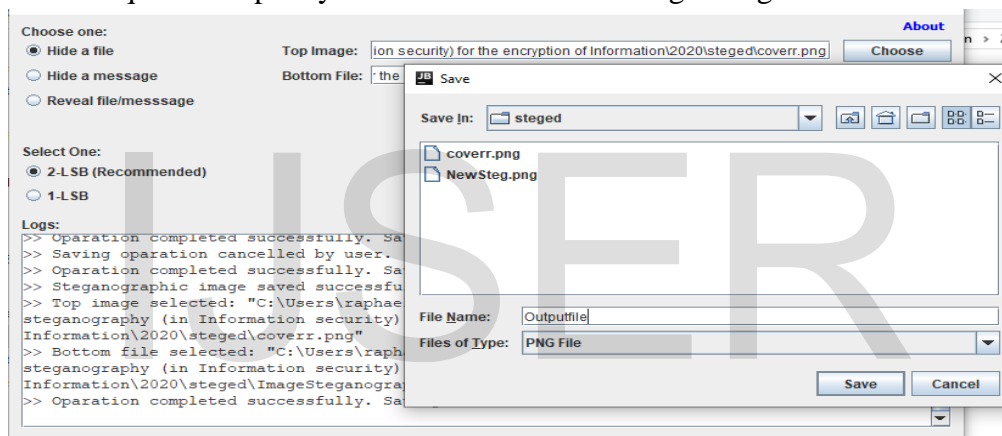


Figure 4.6: Save file

vii. **Decipher the Stego Image.**

This process is used to reverse the initial Cipher process. The procedure is to select the “**Reveal file/message**” option and then choose the output of step vi (stego image) above and insert the password if one is available for the file. Then the user clicks on “Do it” as shown below.

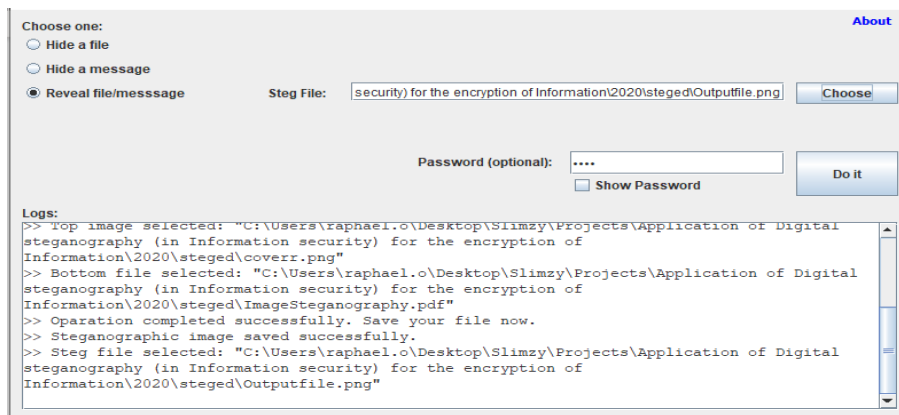


Figure 4.7: Decipher the Stego Image

The user will be asked to specify the location which the deciphered file will be saved on the system as shown below.

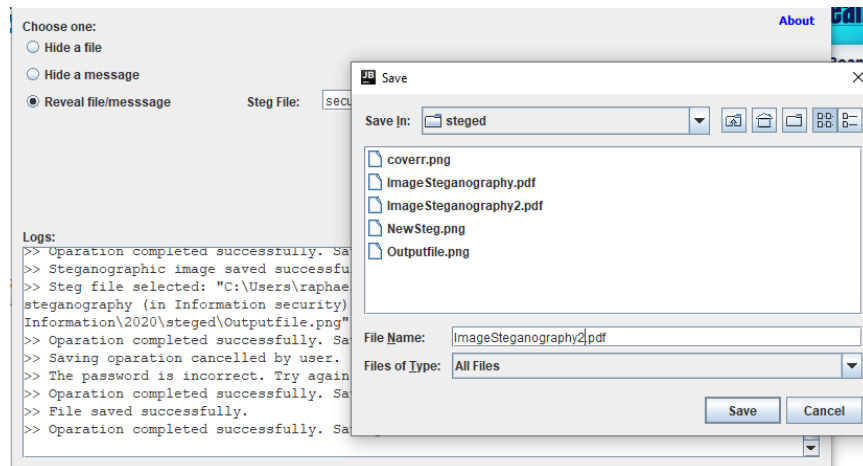


Figure 4.8: Save the deciphered stego image

5.0 CONCLUSION

This research area is very important part when we look at the future of internet security and how to keep data private on open system platforms such as the internet. Steganography is normally used to hide vital information within another file so that when the file arrives at any destination, only the actual recipient who the message was intended for knows that a message exists and can access it. As long as attackers will not resist in finding new ways to perpetuate evil, there is also a constant need for determined researchers who will also continue to contribute to existing security features in order to ensure the security field stays some steps ahead of the attackers.

This research will further be a reference material for future researchers who may be seeking to also contribute to the field of data security and steganography in particular. It presents the concept of digital steganography and particularly addresses the use of this technology in the encoding of various file type asides the popularly known text encoding/hiding. One cannot overemphasize the importance of securing files sent over the internet with the rapid surge in the number of attacks inflicted on innocent users while carrying out their required job functions or business transactions on a daily basis. In this paper, we have discussed various methods that attacks on files being shared over the internet are carried out and proposed a method to hide vital information inside a cover image. This paper mainly focuses on improving the quality of the output image so that it is not distorted and thereby raise the suspicion of a potential attacker about the existence of the vital information concealed therein.

IJSER

REFERENCES

- Alaa. F, Gamal. A, El-Sayed. A, (2016). “Steganography Literature Survey, Classification and Comparative Study”.
- Amirthanjan, Akila. R, Deepika. R, (2015). “A Comparative Analysis of Image Steganography, International Journal of Computer Application”.
- Anupriya. A, Sarita. S, (2018). “A Literature Review on Various Recent Steganography Techniques”.
- Bandyopadhyay, S. K., & Maitra, I. K. (2010). An Alternative approach of steganography using reference image. *arXiv preprint arXiv:1007.1233*.
- Bandyopadhyay S., (2017). “An Alternative Approach of Steganography Using Reference Image. International Journal of Advancements in Technology”.
- Bridgeport. B, Science. C (2017) “A novel video Steganography algorithm in the wavelet Domain Based on the KLT Tracking algorithm and BCH Codes”.
- C. Li, D. Lin, J. Lü and F. Hao (2018), "Cryptanalyzing an Image Encryption Algorithm Based on Autoblocking and Electrocardiography," in IEEE MultiMedia, vol. 25, no. 4, pp. 46-56, 1 Oct.-Dec. 2018, doi: 10.1109/MMUL.2018.2873472.
- Cox, Miller. I, Bloom. M, Fridrich. J, Kalker. J, (2018). “Digital watermarking and Steganography. 4thEd. Elsevier”.
- Hallur. S, Kuri. S, Sudi. S, (2017) “A Robust Digital Watermarking For Gray Scale Image,” International Journal for Technological Research in Engineering, vol. 2, no. 10, pp. 2440–2443, 2017”.
- Lakshman. N, Pedra. G, Ashok. K, (2018). “Different techniques for hiding the text information using text steganography techniques”.
- Mehdi. H, Wahid. A, (2018). “Image steganography in spatial domain: A survey”.
- Moody, G. D., Siponen, M., & Pahnla, S. (2018). Toward a unified model of information security policy compliance. *MIS quarterly*, 42(1).
- Obaida Mohammad, Awad Al-Hazaimah, (2011). “Hiding Data in Images Using New Random Technique Department of Information Technology, AL-BALQA Applied University/Al-Huson University College, Irbid, Al-Huson, 50, Jordan”.
- Ratheeswari, K. (2018). Information communication technology in education. *Journal of Applied and Advanced Research*, 3(1), 45-47.
- Samir Bandyopadhyay, Debnath Bhattacharyya, Debashis Ganguly, Swarnendu Mukherjeet. (2016) “A Tutorial Review on Steganography. Das University of Calcutta”.
- Shawn D, (2017). “An Overview of Steganography at James Madison University”.
- Xingyuan Wang, Le Feng, Hongyu Zhao, (2019). Fast image encryption algorithm based on parallel computing system, *Information Sciences*, Vol. 486, Pages 340-358, ISSN 0020-0255, <https://doi.org/10.1016/j.ins.2019.02.049>. (<https://www.sciencedirect.com/science/article/pii/S002025519301641>)