

A Cost Effective Image Steganography Application for Document Security

Amaobi Uwaleke Michael, Amadi Emmanuel Chukwudi, Nwokonkwo Obi Chukwuemeka

Department of Information Management Technology, Federal University of Technology, Owerri, Nigeria

Email address:

emmanuel.amadi@futo.edu.ng (A. E. Chukwudi)

To cite this article:

Amaobi Uwaleke Michael, Amadi Emmanuel Chukwudi, Nwokonkwo Obi Chukwuemeka. A Cost Effective Image Steganography Application for Document Security. *Management Science and Information Technology*. Vol. 2, No. 2, 2017, pp. 6-13.

doi: 10.11648/j.msit.20170202.11

Received: May 25, 2017; **Accepted:** June 6, 2017; **Published:** July 13, 2017

Abstract: Steganography is the art or practice of concealing a message, image, or file within another message, image, or file. In image steganography, a document is not just encrypted or locked as conventional security techniques could provide, which could leave room for hacking attempts; rather the documents are disguised from their original nature (document hiding behind an image) and as such hacking attempt may not be considered. The image can only be retrieved by individuals with the applications running on their computers. This paper presents an easy to use light weight application that can be used to secure documents within images. This application is developed using C programming language and can be applied in any organization ranging from educational institutions, government agencies financial institutions and private agencies.

Keywords: Encrypted, Steganography, Image, Document, Image

1. Introduction

Any insecure information is prone to be exposed to intruders who are quite alert to reveal the information they got to others, change it to pervert an individual or organization, or exploit it to launch an attack. Various tools and ways have been adopted to protect and secure information especially on the internet. The most commonly use is cryptography mainly due to its simplicity as well as its muddled nature. This method however, is obviously inefficient due to its overt nature of announcing the so-called secured information to the intruders, thereby inviting the intruders to launch attacks on such confidential information. Also, manifold efficacious tools have been set sailed to unveil information locked up using this type of information security tool. To put an end to this unauthorized access of such confidential information, there is a dare need to employ one of the modern information security tools called steganography.

According to Vahid Jessica and Remi [1], Standardized image sources are necessary for development of steganography and steg-analysis. Spatial representation of images can, be very diverse when it comes to the strength and type of noise as well as the complexity of textures.

Statistical properties of pixels can change dramatically after filtering, compression, and resizing. As such the approach to steganographic system design depends on the area of application and the available and familiar tools/technology at the disposal of the developer.

According to Bender et al [2] steganography is a technique of hiding information in digital media. In contrast to cryptography, it is not to keep others from knowing the hidden information but it is to keep others from thinking that the information even exists.

Information hiding is an emerging research area, which encompasses applications such as watermarking, fingerprinting, copyright protection for digital media, and steganography. In watermarking applications for instance, the message contains information such as owner identification and a digital time stamp, which usually applied for copyright protection [2] "Regarding fingerprint, the owner of the data set embeds a serial number that uniquely identifies the user of the data set. This adds to copyright information to make it possible to trace any unauthorized use of the data set back to the user." [3].

Besides cryptography, steganography can be employed to secure information. In steganography, [4] has it that, the message or encrypted message is embedded in a digital host before passing it through the network, thus the existence of

the message is unknown. Besides hiding data for confidentiality, this approach of information hiding can be extended to copyright protection for digital media: audio, video and images.

Steganography becomes more important as more people join the cyberspace revolution. According to Silman [5], "steganography is the art of concealing information in ways that prevents the detection of hidden messages. Steganography includes an array of secret communication methods that hide the message from being seen or discovered."

This work provides a small easy to use application software that can be installed on a computer and used to embed sensitive document within images before sending them over a network link either locally or over the internet. This application can be adopted by any organization ranging from financial institutions, educational institution, government agencies and so on.

2. Statement of Problem

Over the years, information security has been a point of deliberation even as the outstanding tools seem sometimes inept and absolutely inefficient in online data storage and transmission. The most frequently used instrument which is cryptography consists of linguistic or language forms of hidden writing and thus detrimental to a successful information security especially when Internet is involved. One disadvantage of linguistic steganography (cryptography) according to Moerland [4], is that users must equip themselves to have a good knowledge of linguistry. In recent years, everything is tending toward digitization. And with the development of the Internet technology, digital media can be transmitted conveniently over the network. Therefore, messages can be secretly carried by digital media by using the steganography techniques, and then be transmitted through the Internet rapidly.

Steganography is the art of hiding the fact that communication is taking place, by hiding information in other information. Many different carrier file formats can be used, but digital images are the most popular because of their frequency on the internet. For hiding secret information in images, there exists a large variety of steganography techniques. Some are more complex than others and all of them have respective strong and weak points.

Consequently, this security system application will address the problem of information security in a standalone system, online storage and transmission by making confidential information concealment simpler and user friendly.

3. Objective of Study

The goal of steganography is covert communication. So, a fundamental requirement of this steganography system is that the hidden message carried by stego-media should not be sensible to human beings.

The other goal of steganography is to avoid drawing

suspicion to the existence of a hidden message. This approach of information hiding technique has recently become important in a number of application areas.

This research project has following specific objectives:

- i. To present an easy to use security application based on steganography techniques.
- ii. To demonstrate how this application can be used to secure documents

4. Review of Literature

This section reviews key concepts that relates to information security. It further concentrates on the technology behind image steganography.

Cryptography was created as a technique for securing the secrecy of communication and many different methods have been developed to encrypt and decrypt data in order to keep the message secret. Unfortunately it is sometimes not enough to keep the contents of a message secret, it may also be necessary to keep the existence of the message secret. The technique used to implement this, is called steganography.

Steganography is the art and science of invisible communication. This is accomplished through hiding information in other information, thus hiding the existence of the communicated information. The word steganography is derived from the Greek words "*stegos*" meaning "cover" and "*graficT*" meaning "writing" defining it as "covered writing". In image steganography the information is hidden exclusively in images.

4.1. Steganography vs Cryptography

Steganography differs from cryptography in the sense that where cryptography focuses on keeping the contents of a message secret, steganography focuses on keeping the existence of a message secret. Steganography and cryptography are both ways to protect information from unwanted parties but neither technology alone is perfect and can be compromised. Once the presence of hidden information is revealed or even suspected, the purpose of steganography is partly defeated [6]. The strength of steganography can thus be amplified by combining it with cryptography.

4.2. Watermarking and Fingerprinting

Two other technologies that are closely related to steganography are watermarking and fingerprinting as Anderson and Petitcolas [7] rightly put it. These technologies are mainly concerned with the protection of intellectual property, thus the algorithms have different requirements than steganography. These requirements of a good steganographic algorithm will be discussed below. In watermarking all of the instances of an object are "marked" in the same way. The kind of information hidden in objects when using watermarking is usually a signature to signify origin or ownership for the purpose of copyright protection [8].

With fingerprinting on the other hand, different, unique marks

are embedded in distinct copies of the carrier object that are supplied to different customers. This enables the intellectual property owner to identify customers, who break their licensing agreement by supplying the property to third parties [7].

In watermarking and fingerprinting the fact that information is hidden inside the files may be public knowledge - sometimes it may even be visible - while in steganography the imperceptibility of the information is crucial [6]. A successful attack on a steganographic system consists of an adversary observing that there is information hidden inside a file, while a successful attack on a watermarking or fingerprinting system would not be to detect the mark, but to remove it [7].

Research in steganography has mainly been driven by a lack of strength in cryptographic systems. Many governments have created laws to either limit the strength of a cryptographic system or to prohibit it altogether, forcing people to study other methods of secure information transfer [9]. Businesses have also started to realize the potential of steganography in communicating trade secrets or new product information. Accordingly, Artz [10] puts it that, "avoiding communication through well-known channels greatly reduces the risk of information being leaked in transit". Hiding information in a photograph of the company picnic is less suspicious than communicating an encrypted file.

4.3. Steganography Technique

The word steganography is of Greek origin and means "concealed writing" from the Greek words *steganos* (oxsyavo[^]) meaning "covered or protected", and *graphei* (ypacpfi) meaning "writing". The word steganography comes from the Greek "Seganos", which mean covered or secret and "graphy" which mean writing or drawing. Therefore, steganography mean, literally, covered writing. It is the art and science of hiding information such that its presence cannot be detected and a communication is happening. Secret information is encoding in a manner such that the very existence of the information is concealed. Paired with existing communication methods, steganography can be used to carry out hidden exchanges. [5]

The basic model of steganography consists of Carrier, Message and password. Carrier is also known as cover-object, which the message is embedded and serves to hide the presence of the message. Basically, the model for steganography is shown on following figure:

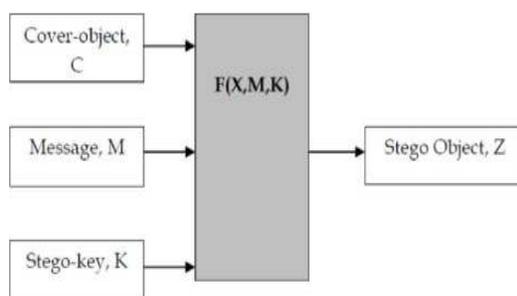


Figure 1. A steganography model.

The Message is the data that the sender wishes to make confidential. It can be a plain text, ciphertext, other image, or anything that can be embedded in a bit stream such as a copyright mark, a covert communication, or a serial number. The Password is known as stego-key, which ensures that only recipient who knows the corresponding decoding key will be able to extract the message from a cover-object. The cover-object with the secretly embedded message is then called the Stego-object. Bender et al [2] stated clearly that, recovering message from a stego-object requires the cover-object itself and a corresponding decoding key if a stego-key was used during the encoding process. The original image may or may not be required in most applications to extract the message.

4.4. Applications of Steganography

Steganographic technologies are very important part of the future of Internet security and privacy on open systems such as the Internet. Steganographic research is primarily driven by the lack of strength in the cryptographic systems on their own and the desire to have complete secrecy in an open-systems environment. Many governments have created laws that either limit the strength of cryptosystems or prohibit them completely. This has been done primarily for fear by law enforcement not to be able to gain intelligence by wiretaps, etc. This unfortunately leaves the majority of the Internet community either with relatively weak and a lot of the times breakable encryption algorithms or none at all.

Civil liberties advocates fight this with the argument that "these limitations are an assault on privacy". This is where Steganography comes in. Steganography can be used to hide important data inside another file so that only the parties intended to get the message even knows a secret message exists. To add multiple layers of security and to help subside the "crypto versus law" problems previously mentioned, it is a good practice to use Cryptography and Steganography together. As mentioned earlier, neither Cryptography nor Steganography are considered "turnkey solutions" to open systems privacy, but using both technologies together can provide a very acceptable amount of privacy for anyone connecting to and communicating over these systems.

The use of steganography can be applied in:

- The educational sector for hiding question papers or important confidential memos.
- In the military for hiding important operational procedures as the case may be
- In the financial sector for hiding financial records and the like
- In government for hiding specialized government documents and a host of other areas of application.

4.5. Steganographic Techniques

Steganography equation is 'Stego-medium = Cover medium + Secret message + Stego key'. The general model of data hiding can be described as follows. The embedded data is the message that one wishes to send secretly. It is usually hidden in an innocuous message referred to as a cover- text or cover-

image or cover-audio as appropriate, producing the stego-text or other stego-object. A stego-key is used to control the hiding process so as to restrict detection and /or recovery of the embedded data to parties who know it [11].

Steganography is classified into 3 categories [12]

- a. Pure steganography where there is no stego-key. It is based on the assumption that no other party is aware of the communication.
- b. Secret key steganography where the stego-key is exchanged prior to communication. This is most susceptible to interception.
- c. Public key steganography where a public key and a private key is used for secure communication.

Steganography methods can be classified mainly into six categories, although in some cases exact classification is not possible [13]

- a. Substitution methods substitute redundant parts of a cover with a secret message (spatial domain).
- b. Transform domain techniques embed secret information in a transform space of the signal (frequency domain)
- c. Spread spectrum techniques adopt ideas from spread spectrum communication.
- d. Statistical methods encode information by changing several statistical properties of a cover and use hypothesis testing in the extraction process.
- e. Distortion techniques store information by signal distortion and measure the deviation from the original cover in the decoding step.
- f. Cover generation methods encode information in the way a cover for secret communication is created..

4.6. Image Steganography

Given the proliferation of digital images, especially on the Internet, and given the large amount of redundant bits present in the digital representation of an image, images are the most popular cover objects for steganography.

In the domain of digital images many different image file formats exist, most of them for specific applications. For these different image file formats, different steganographic algorithms exist. Image steganography techniques can be divided into two groups: those in the Image Domain and those in the Transform Domain [5]. Image - also known as *spatial* - domain techniques embed messages in the intensity of the pixels directly, while for transform - also known as *frequency* - domain, images are first transformed and then the message is embedded in the image. [14].

5. System Design

To discuss about the system design of the new system various variables are put into consideration like the input specification and design, output specification and file design etc. This chapter further elaborates on the variables that make up the new system. Steganography system requires any type of image file and the information or message that is to be hidden. It has two modules encrypt and decrypt.

Microsoft.Net framework prepares a huge amount of tool

and option which helps in simplifying the programming. One of such.Net tools for pictures and images is "auto-converting most types of pictures to BMP format". I used this tool in this software called "Steganography" that is written in C#.Net language. Hence, it is quite easy to use this software to hide in information in any format of pictures without any need of converting its format to BMP. In other words, the software converts the picture on itself.

5.1. Input Specification Design

The basic inputs in the current stego-system consist of the following: Image file and information file.

The format specification of the image file is bitmap while the information file can be any type, viz:.doc, .docx, .pdf, .xls etc.

The size of the image file determines the size of information file to hide in the image. The formula connecting these two parameters is given below:

$$S = (8.0 * ((height * (width / 3) * 3) / 3 - 1)) / 1024;$$

Note that: Width = width of image file,

Height = height of image file

S = maximum size of information that can be embedded by the image

5.2. Algorithm Used

The algorithm used for Encryption and Decryption in this application is provided using several layers in place of using only LSB layer of image. Writing data starts from last layer (8st or LSB layer); because significant of this layer is least and every upper layer has doubled significant from its down layer. So every step we go to upper layer image quality decreases and image retouching transpires.

The rationale behind the use of the Least Significant Bit (LSB) algorithm is because insertion is easy and requires simple approach to embedding information in a cover image.

In other words, by using LSB algorithm, storing 3 bits in each pixel of the image is made possible. An 800 x 600 pixel image, can thus store a total amount of 1,440,000 bits or 180,000 bytes of embedded data. For example a grid for 3 pixels of a 24-bit image can be as follows:

```
(00101101 00011100 11011100) (10100110 11000100
00001100) (11010010 10101101 01100011)
```

When the number 200, which binary representation is 11001000, is embedded into the least significant bits of this part of the image, the resulting grid is as follows:

```
(00101101 00011101 11011100)
(10100110 11000101 00001100)
(11010010 10101100 01100011)
```

Although the number was embedded into the first 8 bytes of the grid, only the 3 underlined bits needed to be changed according to the embedded message. On average, only half of the bits in an image will need to be modified to hide a secret message using the maximum cover size. Since there are 256 possible intensities of each primary colour, changing the LSB of a pixel results in small changes in the intensity of the colours. These changes cannot be perceived by the human eye - thus the message is successfully hidden. With a well-chosen

image, one can even hide the message in the least as well as second to least significant bit and still not see the difference.

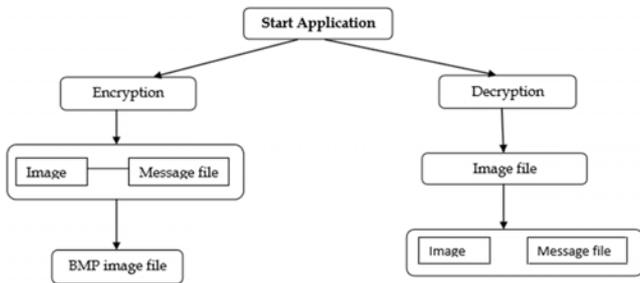


Figure 2. Block Diagram of the Image Steganography system.

5.3. Justification of Programming Language Used

C# is a multi-paradigm programming language encompassing strong typing, imperative, declarative, functional, generic, object-oriented (class-based), and component-oriented programming disciplines. It was developed by Microsoft within its.NET initiative and later approved as a standard by Ecma (ECMA-334) and ISO (ISO/IEC 23270: 2006). C# is one of the programming languages designed for the Common Language

Infrastructure (CLI).

The rationale behind the choice of C# in the design of this system is conspicuously highlighted below.

- a. C# language is a simple, modern, general-purpose, object-oriented programming language. C# attempts to simplify the syntax to be more consistent and more logical in the design of the system.
- b. It provides software robustness, durability, and makes programmer productivity easy.
- c. It enhances source code portability, as is programmer portability.
- d. C# is suitable for writing applications for both hosted and embedded systems, ranging from the very large that use sophisticated operating systems, down to the very small having dedicated functions.
- e. C# applications are more economical with regard to memory and processing power requirements. C# removes memory management issues from the developer by using.NET's garbage collection scheme. Items no longer referenced are marked for garbage collection, and the Framework can reclaim this memory as needed.

5.4. Flow Chart of the New System Model

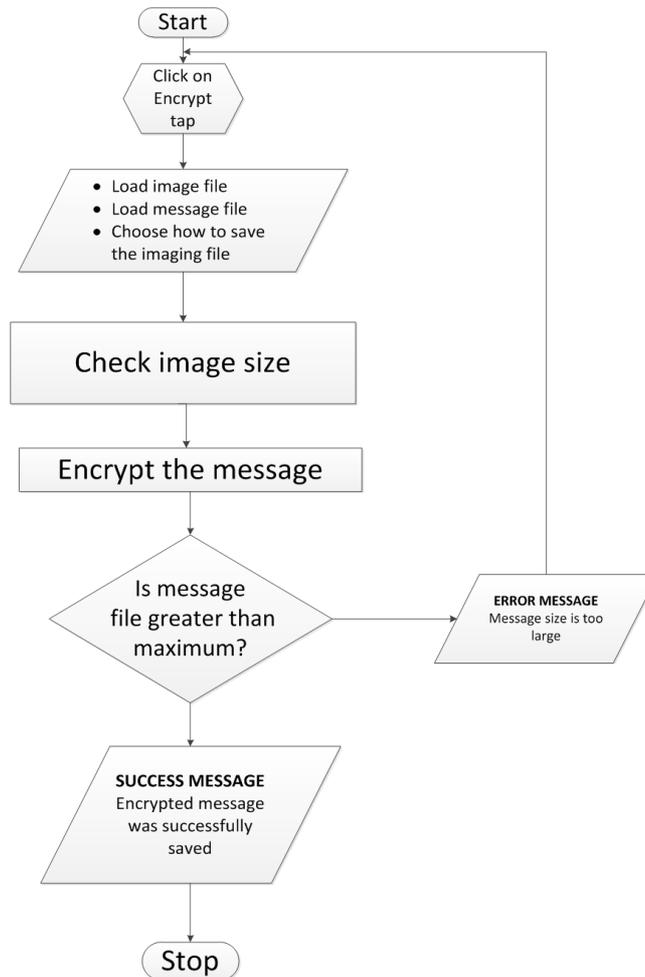


Figure 3. System flow chart.

5.5. The System Model and Output Screens and Implementation

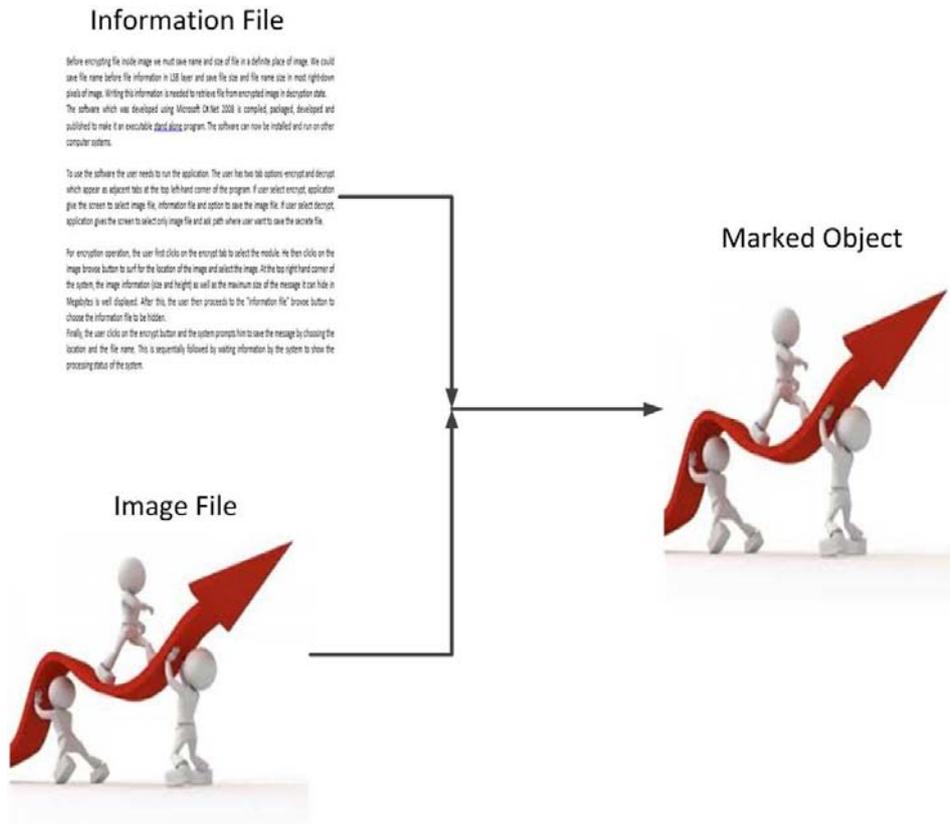


Figure 4. Encryption process model.

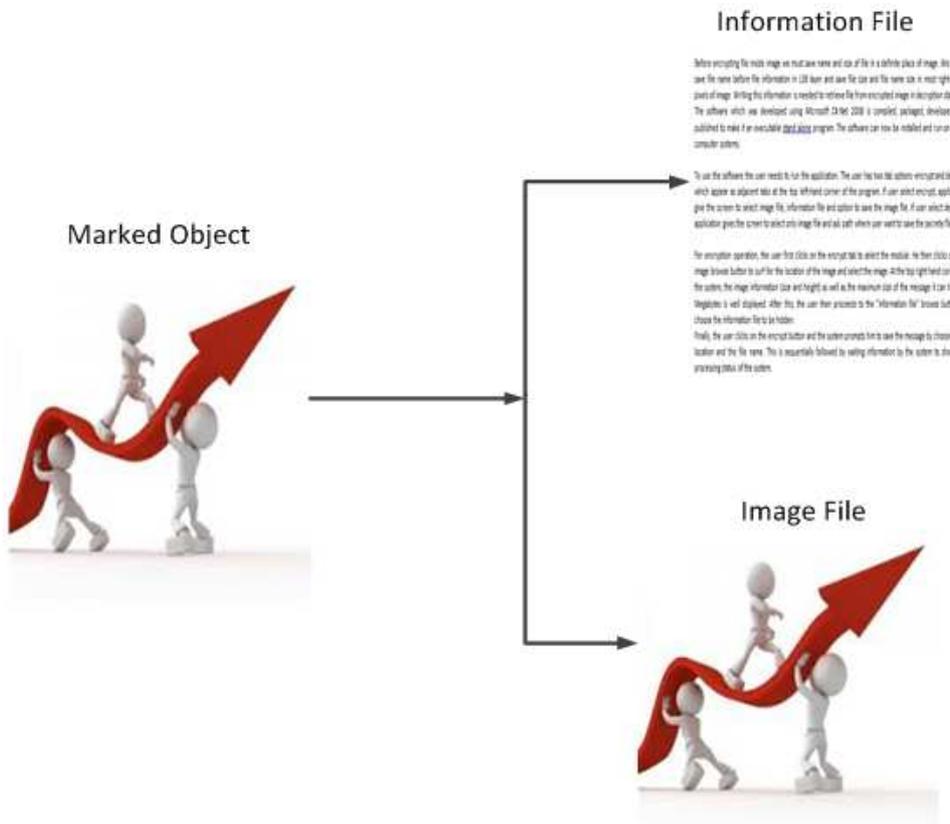


Figure 5. Decryption process model.

In figure 5, a document is hidden i.e. encrypted inside an image to form a “marked object”, while the marked object in figure 6 is decrypted to generate the original document encrypted. This process describes the steganography process described in the paper.



Figure 6. The steganography system image select screen.



Figure 7. The steganography system after selecting the image and file.

The system is made up of two modules: encrypt and decrypt modules. The encrypt module is used to hide information into the image; no one can see that information or file. This module requires any type of image and message and gives the only one image file in destination. The decrypt module is used to get the hidden information in an image file. It takes the image file as an output, and gives two files at destination folder, one is the same image file and the other is the message file that is hidden in it.

Before encrypting file inside image we must save name and size of file in a definite place of image. We could save file name before file information in LSB layer and save file size and file name size in most right-down pixels of image. Writing this information is needed to retrieve file from encrypted image in decryption state.

The software which was developed using Microsoft C#.Net 2008 is compiled, packaged, developed and published to make it an executable stand alone program. The software can now be installed and run on other computer systems.

To use the software the user needs to run the application. The user has two tab options -encrypt and decrypt which appear as adjacent tabs at the top left-hand corner of the program. If user select encrypt, application give the screen to select image file, information file and option to save the image file. If user select decrypt, application gives the screen to select only image file and ask path where user want to save the secrete file.

For encryption operation, the user first clicks on the encrypt tab to select the module. He then clicks on the image browse button to surf for the location of the image and select the image. At the top right hand corner of the system, the image information (size and height) as well as the maximum size of the message it can hide in Megabytes is well displayed. After this, the user then proceeds to the "information file" browse button to choose the information file to be hidden.

Finally, the user clicks on the encrypt button and the system prompts him to save the message by choosing the location and the file name. This is sequentially followed by waiting information by the system to show the processing status of the system.

On the other hand, to decrypt an already encrypted message using the system, the user first selects the decrypt tab followed by the image browse button to select the message to be decrypted. At the tail end of the process, the user selects the "decrypt button" and where to save the decrypted message.

6. Conclusion and Recommendation

6.1. Summary

Although only some of the main image steganographic techniques were discussed in this paper, one can see that there exists a large selection of approaches to hiding information in images. All the major image file formats have different methods of hiding messages, with different strong and weak points respectively. Where one technique lacks in payload capacity, the other lacks in robustness. For example, the patchwork approach has a very high level of robustness against most type of attacks, but can hide only a very small amount of information.

Least significant bit (LSB) in both BMP and GIF makes up for this, but both approaches result in suspicious files that increase the probability of detection when in the presence of a warden.

Thus for an agent to decide on which steganographic algorithm to use, he would have to decide on the type of application he want to use the algorithm for and if he is willing to compromise on some features to ensure the security of others.

In addition, it has been proved beyond reasonable doubt that, stego-system has quite a multitudinous applications both in individual transactions and business dealings. It complements the existing system to leave up legal band and

to reinforce the present level of security.

6.2. Recommendations

Steganography is seen as a high-level type of encryption; hence, it can be used in information security within institutions as its use will result in a mechanism to implement two of the five key pillars of information security, namely confidentiality and integrity. Here, the confidentiality of the hidden message is protected due to it being unrecognisable in its hidden and encrypted form both in the place of storage and during transmission while the encrypting of the concealed message protects the integrity of the data.

Besides this, more research is recommended to be done in the area of stego-key provision to ensure absolute lock of the information being transmitted in this stego-system. Again, applications of steganography are wide ranging, and are indeed valuable if used in the correct manner. Therefore, I would like to recommend the use of this new technology in business and individual transactions to reduce cost and enhance the net revenue of such business or individual as the case may be. Since neither Cryptography nor Steganography are considered "turnkey solutions" to open systems privacy, the use of both technologies together to provide a very acceptable amount of privacy for anyone connecting to and communicating over these systems is required.

Finally, since this security technology may be misused and abused, resulting in disastrous consequences, Use of official instrument to control the transmission of embedded information through Steganography to check the use of such tool in concocting criminal and inhumane plots in the society.

References

- [1] Sedighi, V., Fridrich, J., & Cogranne, R. (2016). Toss that BOSSbase, Alice!. *Electronic Imaging*, 2016 (8), 1-9.
- [2] Bender, W., Gruhl, D., Morimoto, N. & Lu, A., (1996). Techniques for data hiding. *IBM Systems Journal*, 35 (2).
- [3] Dunbar, B. (2002). Steganographic techniques and their use in an Open-Systems environment. SANS Institute, January.
- [4] Moerland, T.(2001). Steganography and Steganalysis. Leiden Institute of Advanced Computing Science. Accessed September 12, 2012. Available from www.liacs.nl/home/tmoerl/privtech.pdf
- [5] Silman, J.,(2001). Steganography and Steganalysis: An Overview. SANS Institute.
- [6] Wang, H & Wang, S. (2004). Cyber warfare: Steganography vs. Steganalysis. *Communications of the ACM*, 47 (10) October.
- [7] Anderson, R. J. & Petitcolas, F. A. (1998). On the limits of steganography. *IEEE Journal of selected Areas in Communications*, (May): 22.
- [8] Marvel, L. M., Boncelet Jr., C. G. & Retter, C. (1999). Spread Spectrum Steganography. *IEEE Transactions on image processing*, 8 (08).
- [9] Mondal, Saikat, Rameswar Debnath, and Borun Kumar Mondal. "An improved color image steganography technique in spatial domain." *Electrical and Computer Engineering (ICECE), 2016 9th International Conference on*. IEEE, 2016.
- [10] Artz, D. (2001). Digital Steganography: Hiding Data within Data. *IEEE Internet Computing Journal*, (June).
- [11] Fabien A. P. Petitcolas, Ross J. Anderson and Markus G. Kuhn, (1999) "Information Hiding – A Survey", *Proceedings of the IEEE*, special issue on protection of multimedia content, pp. 1062-1078.
- [12] C. P. Sumathi, T. Santanam and G. Umamaheswari (2013) A Study of Various Steganographic Techniques Used for Information Hiding. *International Journal of Computer Science & Engineering Survey (IJCSES)* Vol. 4, No. 6.
- [13] Stefan Katzenbeiser & Fabien A. P. Petitcolas (1999), *Information Hiding Techniques for Steganography and Digital Watermarking*, Artech House, Computer Security series, Boston, London.
- [14] Lee, Y. K. & Chen, L. H. (2000). High capacity image steganographic model. *Visual Image Signal Processing*, 147 (03), June. Maes (Ed.).