Efficient Steganography on Video File using Discrete Cosine Transform Method (DCTM)

Ononiwu R. N.
Department of Computer Science,
University of Port-Harcourt
Port-Harcourt, Nigeria

Okengwu U. A.
Department of Computer Science,
University of Port-Harcourt
Port-Harcourt, Nigeria

ABSTRACT
The desire to keep sensitive information and communications secret for many decades, and potential generations have been paramount. In this era of digital media and internet communications, the needs seem more pressing as lots of information and people's conversations are been tracked, stolen, manipulated and pirated. Steganography is the art of hiding a message, a covert substance (stego) so that only the sender and the intended recipient know that there is a message contained in the substance or the stego. It can also be stated to be the fine art of hiding information. This means that hiding message in a carrier file enables the deniability of the existence of that message. This work provides an overview of steganography, designs software to illustrate specific steganography methods and recent developments in the field. The methodology used to develop this work was object-oriented analysis and design methodology (OOAD), and the programming language used is Java programming language which is based on classes and objects as it took care of compatibility over multiple operating systems. The work was done by embedding text file in a video frame (video file) in such a manner that the video file does not lose its functionality using Discrete Cosine Transform Method (DCTM). This method strives for high data security to a hacker's inability to detect the existence of a hidden message or information.

Keywords
Steganography, Cryptography, Steganalysis

1. INTRODUCTION
The escalation of computer technology and the internet have made breakthrough in the existence of data and its communication. This has opened a new way of implementing steganography to ensure secure data transfer. A phrase "information is power" is significant in computer systems. Through "information" a lot of data are gathered. Important data should be secured from unauthorized access and users. One of the solutions which enable us to store data in a secure way and check their authenticity and confidentiality is the use of cryptographic methods, which is known as cryptography.

There is no such thing as being anonymous or hiding identity on the internet because of the extensive technological developments. We cannot avoid expressing ourselves on the internet, but the only thing we can do is to take care of our data confidentiality. Contrary to the opinions about ever-present hackers waiting to steal information, we can still secure our data, and prevent them from listening to our internet conversations. It does not matter if we are talking about our professional issues or private ones, we cannot let unauthorized persons gain access to our information.

There is no doubt about data been the most precious thing which is stored in our computers. In the case of stealing, we can easily replace the hardware, but data usually cannot be replaced easily. It is also not such a serious problem if a device is the target of an attack. It is much worse if an attacker tries to get access to data. In times of common mobility, it is easy to lose a laptop in which important business e-mails, credit cards, numbers, or official papers are stored. In such situation, proper access rights, good anti-virus program or firewall, appropriate system configuration will render the stolen systems useless and grant us safety. In the case of digital information, data encryption and steganography will be our last resort, as encrypted data is often useless for the third parties. The use of proper encrypting tools guarantee the cipher will be practically unbreakable.

Steganography is the fine art of hiding information or practice of concealing a file, image, video or message within another file, in ways that prevent the detection of the hidden messages. The word "steganography" is of Greek origin, and it means "covered writing" or "concealed writing". Steganography is changing the digital media in a way that only the sender and the intended recipient is able to detect the message sent through it. On the other hand, steganalysis is the science of detecting hidden message. The objective of steganalysis is to break steganographic system and such condition is met if an algorithm can judge or show whether a given image contains secret messages. To reduce the possibility of attack, security needs to be kept secret, that is, invisible security. The important data can be inserted into multimedia documents in a way that cannot be spotted, that is, imperceptible (invisible) insertion of information into multimedia data.

2. STATEMENT OF THE PROBLEM
Steganography became inevitably more important as many people join the cyberspace revolution. Lots of people are destructive and dangerous using hijacked data and intercepted multimedia information. However, there are classic methods of securing communications base on cryptography, which encrypts plain text to generate cipher text, yet the transmission of cipher text may easily arouse attackers' suspicion. In order to make up for the shortcomings of cryptographic techniques, steganography has been developed as a new covert communication means in recent years. Steganography transfers message secretly by embedding the message into a cover medium with the use of information hiding techniques. The goal of steganography is to avoid drawing suspicion to the existence of a hidden message.

3. AIM AND OBJECTIVES OF THE STUDY
The aim of this work is to design steganography system with video file.

The objectives include:

i. to develop a model for hiding and revealing messages using discrete cosine transform method (DCTM)

ii to design a software that will use the DCTM model to perform steganography.

This program will hide and retrieve data in a digital object (video frame) using a mathematical method known as Discrete Cosine Transform method (DCTM).

The significance of this study is the implementation of Discrete Cosine Transform Method. This is achieved, when the disclosure of any hidden message remains difficult, as the hacker (attacker) cannot provide the secret key (password), thereby hindered with the problem of decoding the hidden message since the sequence of the images within the video file is unknown to the attacker.

The scope of the study is the implementation of the steganographic tools for hiding information which includes texts, image files, and the exploitation of the Discrete Cosine Transform Method (DCTM).

The limitations of the study are the inability of some computer science experts to share their knowledge and views on the subject matter. Some suspected hackers interviewed were unable to state clearly, the reasons why they are mischievous on data and information on-line.

4. STEGANOGRAPHY AND INFORMATION HIDING

The existing systems lack good user interface, non provision of choosing the secret key (password) and more encode-decode time consumption. There are numerous steganography programs available. A few of them are excellent in every respect; unfortunately, most of them lack usable interfaces, some contain too many bugs, and unavailability of a program for other operating systems. This application will take into account these shortcomings, and since it was written in Java programming language, compatibility over multiple operating systems, even different hardware platforms would not be an issue. The idea behind this design is to provide an easy and efficient method for hiding data from hackers and send to the destinations securely. This system is based on video steganography for hiding data in the video image, retrieving the hidden data from the video image using Discrete Cosine Transform method (DCTM). The design looks at a specific class of commonly used image, based on steganographic techniques. Under what condition can an observer distinguish between stego images (images which carry secret messages) and cover images (images that do not carry secret message). Figure 1 shows two video images, the "carrier image" which is the image without any message, and the "stego image", which is the image that contains hidden message. It is not viable to identify the difference between the original picture (the carrier image) and the stego image (the picture that has a secret message embedded in it).

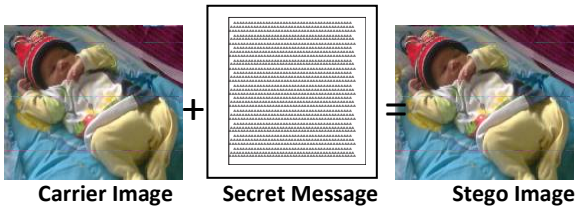

Figure 1: Steganography using Video Image

Steganography is the art of hiding and transmitting data through apparently harmless and safe carriers in an effort to conceal the existence of the data. Steganography has its place in security, it is not intended to replace cryptography rather steganography supplements cryptography. Consequently, cryptography and steganography are two important branches of information security. However cryptography and steganography are often bridged by the use of steganalysis, because of some weak algorithms often used.

Hiding a message with Steganographic methods reduces the chance of a message being detected. If the message is also encrypted then it provides another layer of protection (Johnson, 2006).

Therefore, some steganographic methods combine traditional cryptography with steganography, the sender encrypts the secret message prior to the overall communication process, as it is more difficult for an attacker to detect embedded cipher text in a cover (Sellar, 2003).

In the field of steganography, some terminologies have been developed. The adjectives 'cover', 'embedded', and 'stego' were defined at the information hiding workshop held in Cambridge, England. The term "cover" refers to description of the original innocent image, data, audio, video, and so on. The term 'stego' refers to the same image when a message or information is embedded ("embedded" - an adjective that describes the action performed) in it. Steganography is not a new science. It dates back to ancient times. It has been used through the ages by ordinary people, spies, rulers, government, and armies (Sellars, 2002).

The stories about steganography has been pleasant, for example, the story of how ancient Greece used some methods of hiding messages in the tummy of a hare (a kind of rabbits), pigeons to conceal and deliver messages safely and securely to the destination, also the use of invisible ink.

Another ingenious and clever method was to shave the hairs on the head of a messenger and tattooed a message or image on the messenger's head, after allowing new hairs to grow and cover the tattooed message, the message would be undetected until the new grown hair is shaved off the messenger's head again.

Also, the concept of steganography was practiced thousand years ago when the Greek were used to send secret information by writing in wax-covered tablets; wax was first scraped off a tablet, the secret message was written on the tablet, and then the tablet was covered again with the wax. While the Egyptians used illustrations to conceal message (Davern, 2002).

5. CRYPTOGRAPHY AND STEGANOGRAPHY

Since the advent of computers, there has been a vast dissemination of information, some of which needs to be kept private, secret and confidential, some of which does not. Information may be hidden in two basic ways via cryptography and steganography.

The methods of cryptography does not conceal the presence of secret information but renders it unintelligible and incomprehensible to the outsider by various transformations of the information that is to be put into secret form, while the methods of steganography conceals the very existence of the secret information. (Al-Dieimy, 2002; Dorothy, 2000).

Some operation for data transformation

a. Cryptography with Block Cipher: In cryptography, a block cipher is a symmetric key cipher which operates on fixed-length groups of bits, termed blocks, with an unvarying transformation.

When encrypting, a block cipher might take, for example, a 128-bit block of plain text as input, and outputs a corresponding 128-bit block of cipher text. The exact transformation is controlled using a second input — the secret key.

b.Decryption: Decryption is similar, the decryption algorithm takes in this example, a 128-bit block of cipher text together with the secret key, and yields the original 128-bit block of plain text. To encrypt messages longer than the block size (128 bits in the above example), a mode of operation is used. Block ciphers can be contrasted with stream ciphers.

c. Stream Cipher: A stream cipher operates on individual digits, one at a time and the transformation varies during the encryption. The distinction between the two types is not always clear-cut, a block cipher, when used in certain modes of operation, acts effectively as a stream cipher, (Johnson, 2005).

A highly influential block cipher design is the Data Encryption Standard (DES).

6. GENERAL STEGANOGRAPHY SYSTEMS

In a general steganography system, it is assumed that the sender wishes to send a message to a receiver via steganographic transmission. The sender starts with a cover message, which is the medium where the embedded message will be hidden. Note that the hidden message is called the "the embedded message". This is an input to the steganographic system. A steganographic algorithm combines the cover massage with the embedded message to form a steganographic message (stego message). The algorithm may, or may not, use a steganographic key (stego key), which is additional secret data that may be needed in the hidden process. The same key (or related one) is usually needed to extract the embedded massage again. The output of the steganographic algorithm is the steganographic message (stego).

The cover massage and stego must be of the same data type, but the embedded message may be of different data type. The receiver reverses the embedding process to extract the embedded message (Avedissian, 2005).

7. METHODOLOGY

The concept of steganography came into existence many years ago. The use of invisible ink during the world war II, the primitive method of drawing tattoo to send secret messages, including the ancient methods of the Greeks sending secret information, using wax-covered table. In the past years, numerous steganographic techniques that embed hidden messages have been proposed. Also, cryptography had often been used in other to protect digital information without a good encryption standard and algorithm, thereby making the digital information vulnerable to the man in the middle to attack. The proposed system is steganography based on video files for hiding data in the video image, retrieving the hidden data from the video image using Discrete Cosine Transform method (DCTM) on an effective platform. This design looks at a specific class of widely used images based on steganographic techniques, under what conditions can an observer distinguish between steganographic images ("stego image" - images which carry a secret message) and cover images (images that do not carry a secret message). Figure 2 shows two video images, one - the cover image, also known as "Carrier Image" which is the image without any secret message and the other image labeled "Stego Image" which contains the secret message. It is not viable to identify the difference between the original video image and the Stego video image.

Figure 2: Steganography using Video Image

Steganography is not an alternative to cryptography. Steganography supplements cryptography. While cryptography provides privacy, steganography is intended to provide secrecy. In other words, cryptography works to mask the content of a message while steganography works to mask the very existence of the message.

Steganography is changing the digital media in a way that only the sender and the intended recipient is able to detect the message sent through it. The following formula provides a generic description of the pieces of the steganographic process; Cover_medium + hidden data + stego_key = stego_medium.

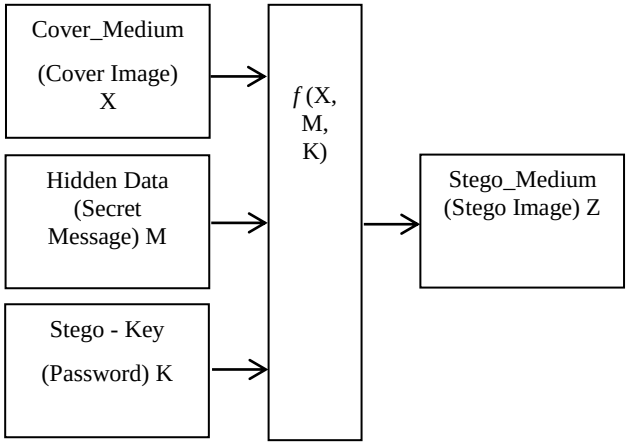

Figure 3: Basic Steganography Model

In this context, the cover_medium (cover image, X) is also known as Carrier, in which the message is embedded, it is used to hide the presence of the message, which may be encrypted. Secret Message, also known as hidden data in this context is the data that the sender wishes to remain confidential. It can be plain text, cipher text, other image, or anything that can be embedded in a bit stream such as a copyright mark, a covert communication, or a serial number. Stego Key is also known as the password, it ensures that only recipients who know the corresponding decoding key (password) will be able to extract the secret message from the stego_medium, which is the resultant file. The stego_medium also known as stego image Z, will of course be the same data type like the cover_medium.

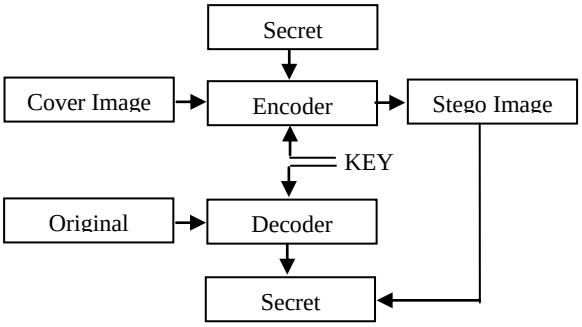

Figure 4: Elements of Steganography

1. The proposed system will provide a usable and good looking Graphical User Interface (GUI) for the system.

2. The system will be interactive and user friendly, it will also show the ability to operate the steganography machine with no prior training and consultation of any help files.

The methodology used to develop this work is object-oriented analysis and design methodology (OOAD) which is based on classes and objects. In this environment, software is the collection of discrete objects that encapsulate their data as well as the functionality to model real world "objects".

A class is a pattern or templete for creating multiple objects with similar features.

An object defines the variables and methods of the classes. Objects declare the capabilities of the classes, however before these capabilities can be used, an object must be created.

Therefore, an object is an instantiation of a class. This means that an object is composed of the memory allocated for the member variables of the class.

Each object has attributes and methods (each object has its own set of member variable). The following occurs when a new object is created:

(i) A new keyword is used to create an instance of a class.

(ii) Memory is physically allocated for the new instance of the class.

(iii) Any static initialize is executed.

(iv) A constructor is called to do initialization.

(v) A reference to the object is returned.

8. METHOD OF CONCEALING DATA IN VIDEO IMAGE:

Discrete Cosine Transform Method (DCTM):

i. Discrete Cosine Transform is a technique widely used in image/video file compression.

ii. Discrete Cosine Transform coefficients are used for FLY, AVI file compression, it is the heart of JPEG and MPEG file formats.

iii. Discrete Cosine Transform Method separates image into parts of differing frequencies. It transforms a signal or image from the spatial domain to the frequency domain.

iv. Discrete Cosine Transform Method can separate the image into high, middle and low frequency components.

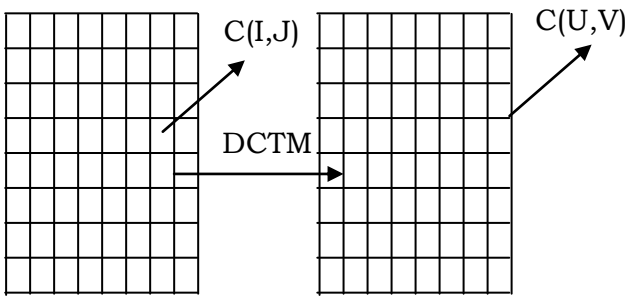

Figure 5: Discrete Cosine Transform of an Image

The general equation for a 1D (N data items) DCTM is defined by the following equation:

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) cos\left(\frac{(2x+1)u\pi}{2N}\right)$$

for u = 0, 1, 2, ..., N-1.

The general equation for a 2D (N by M image) DCTM is defined by the following equation:

$$C(u,v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) cos\left(\frac{(2x+1)u\pi}{2N}\right) cos\left(\frac{(2y+1)v\pi}{2N}\right)$$

For u, v = 0, 1, 2, . . . , N-1.

Source: International Journal of Engineering and Innovative Technology (IJEIT) Volume 1, Issue 4, April 2012.

For u, v = 0, 1, 2, . . . , N-1. Here, the input image is of size N x M. C(i, j) is the intensity of the pixel in row i and column j; C(u,v) is the Discrete Cosine Transform coefficient in row u and column v of the Discrete Cosine Transform matrix. Signal energy lies at low frequency in image; it appears in the upper left corner of the Discrete Cosine Transform. Compression can be achieved since the lower right values represent higher frequencies, and generally small enough to be neglected with little visible distortion. Discrete Cosine Transform Method (DCTM) is used in steganography as; Image is broken into 8×8 blocks of pixels, working from left to right, top to bottom, the Discrete Cosine Transform Method (DCTM) is applied to each block. Each block is compressed using a process known as quantization, which retains the most important frequencies of the image and discards the less important frequencies. The complete array of compressed blocks of the image is stored in drastically reduced amount of space.

When one wants to see the stored image again, it can be done through the process of decompression which uses Inverse Discrete Cosine Transform Method (IDCTM) for the purpose.

9. USE CASE DIAGRAM

A use case describes a sequence of actions that provide something of measurable value to an actor and is drawn as a horizontal ellipse. An actor here is a person, organization, or external system that plays a role in one or more interactions with your system.

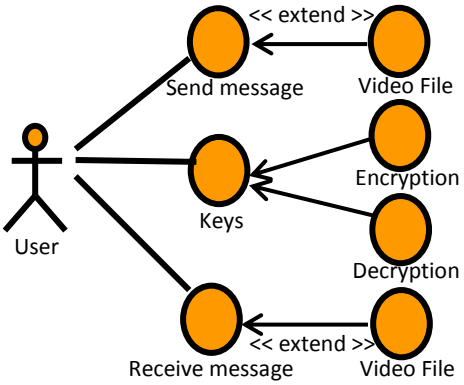

Figure 6: Use Case Diagram

10. ARCHITECTURE OF THE PROPOSED SYSTEM

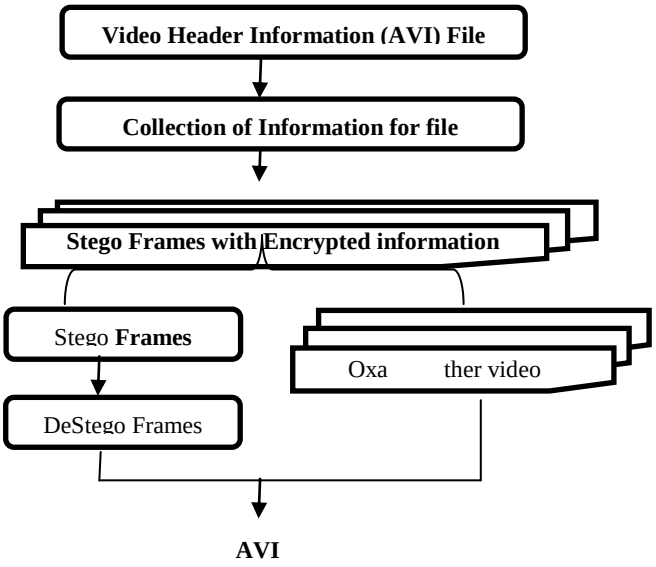

Figure 7: An Architecture of the Proposed System

Architecture

In figure 8 above, which depicts the architectural design of the system, the video steganography software performs the process of concealing and revealing information in the following modules.

1. Video Header Information: Is used to check the file format
2. File Handling: Does the conversion of text files to binary values
3. 3.Encryption: Is a process of hiding message against a third party - Encode a key (password)
4. Steganography: Is the process of concealing data (keeping the hidden message secret)
5. DeSteganography: Is the process of reveal the original data
6. Decryption: Is a process of decoding a hidden message (using same password encoded)
7. Graphical User Interface: Is a wizard that coordinates the relationship between the system, the users and the environs.

11. ALGORITHMS FOR VIDEO STEGANOGRAPHY USING DCTM

Algorithm to Conceal Text Message

Step 1: Read cover image.

Step 2: Read secret message and convert it in binary.

Step 3: The cover image is broken into 8×8 block of pixels.

Step 4: Work from left to right, top to bottom subtract 128 in each block of pixels.

Step 5: Discrete Cosine Transform is applied to each block.

Step 6: Each block is compressed through quantization table.

Step 7: Calculate Least Significant Bits of each Discrete Cosine Transform coefficient and replace with each bit of secret message.

Step 8: Write stego image.

Algorithm to Retrieve Text Message:-

Step 1: Read stego image.

Step 2: Stego image is broken into 8×8 block of pixels.

Step 3: Work from left to right, top to bottom subtract 128 in each block of pixels.

Step 4: Discrete Cosine Transform is applied to each block.

Step 5: Each block is compressed through quantization table.

Step 6: Calculate Least Significant Bits of each Discrete Cosine Transform coefficient.

Step 7: Retrieve and convert each 8 bits into character.

12. RESULT

The result of this work is the ability to design and develop software that can confidentially keep information secret through the use of video frames. This process of hiding data in a video frame involves encoding and decoding processes with the use of a mathematical model known as Discrete Cosine Transform (DCT) Method.

Below are the displays of some of the results of the work.

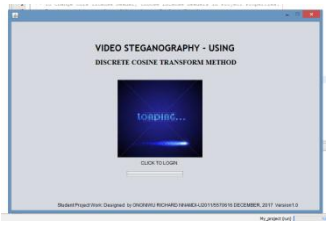

Figure 8: Home page

This page is the first output of the programme, it displays the information of what the system is all about and the methods used in manipulating the system or machine.

The programme is interactive as you follow the instructions and responses.

Figure 9

This is another output of the programme, in this particular environment the user is given the opportunity to select video file where the message will be hidden, type in the secret message, select ENCODE, then enter the password (secret key) and finally click on OK, at this point the message will be compressed and embedded in the video frame of the video image selected where the message will be kept secret.

Figure 10

In this output, the user has selected the video file that will contain the secret message, typed the secret message to be hidden, clicked on encode button and is suppose to give the file a serial number and create a secret key or password through which the file will be locked against the third party or unauthorized users.

Figure 11

The above is an output message to show and confirm that a file has been created and hidden. It also indicates or marks the end of the process of encoding.

Figure 12

This result of the programme marks the beginning of the process of decoding. Here the user is given the opportunity to click on video browser and select the video file that contains the video image where the secret message was embedded and hidden then click on DECODE.

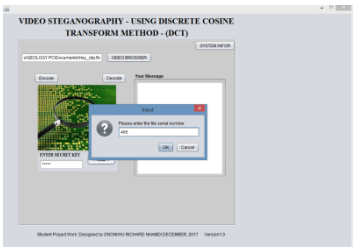

Figure 13

This output is also part of the process of decoding where the user or the recipient is provided with opportunity to enter the secret key or password used to lock up the secret message during the process of encoding.

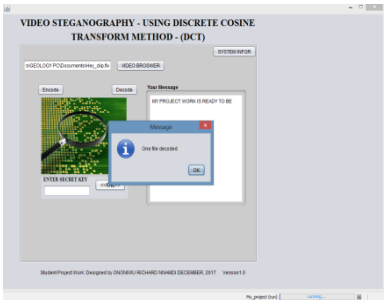

Figure 14

This is the final output or result of the work where the hidden message is displayed on the screen for the recipient's consumption.

13. CONCLUSION

Video frames are created in numbers and a secret message is encoded (hidden) within one of the video frames. A secret key (password) is generated using the proposed software developed (video steganography), the password locks up every access to retrieve or disclose the hidden message to any third party. In summary, no matter the level of attack, no hacker can penetrate or break this system, except the receiver wishes to decode the secret message with the use of the same software developed and the provision of the secret key (password) originally generated during the process of encoding.

14. REFERENCES

[1] Fridrich, J., Goljan, M., and Du R., (2001) “Detecting LSB Steganography in Color and Grayscale Images,” IEEE Trans Multimedia Special Issue on Security, vol 8, Issue 4, pp. 22–28.

[2] Mamta Juneja, Parvinder S. Sandhu, and Ekta Walia, (2009) "Application of LSB Based Steganographic Technique for 8-bit Color Images”, WASET. https://zenodo.org/record/1086191#.Xk_Hf_lKjIU

[3] Mazdak Zamani, Azizah A. Manaf, Rabiah B. Ahmad, Akran M. Zeki and Shahidan Abdullah, (2009) “A Genetic- Algorithm-Based Approach for Audio Steganography” International Journal of Computer, Electrical, Automation, Control and Information Engineering, WASET. vol 3, no 6, https://pdfs.semanticscholar.org/7126/6c91b0806c851238ee328af2806b852a3723.pdf

[4] Neil F. Johnson and Jajodia S., (1998) Exploring Steganography. Seeing the Unseen, IEEE Computer, vol. 31, issue 2, pp 26 - 34. https://ieeexplore.ieee.org/document/4655281/keywords#keywords

[5] Niels Provos and Peter Honeyman, (2003), "Hide and Seek: An Introduction to Steganography", IEEE Security and Privacy Magazine. vol 1, issue 3, pp 32 - 44

[6] Petitcolas, F.A.P., Anderson, R. J., and Kuhn, M. G., (1999) "Information Hiding - A Survey," Proceedings of the IEEE. Vol 87, issue 7, pp 1062 – 1078.

[7] Shannon, C. E. A (1948) Mathematical Theory of Communication; Bell System Technical Journal, vol. 27, pp 379 - 423

[8] Simmons, G. J., (2010) "The Prisoners' Problem and The Subliminal Channel, in Advances in Cryptology"; Proceedings of Crypto 83 (D. Chaum, ed.), Plenum Press, vol 12.9, pp 51-67

[9] Sutaone, M.S., Khandare M. V., (2008) "Image based Steganography using LSB insertion technique", IET. https://ieeexplore.ieee.org/document/4470096/authors#authors

[10] Andreas Westfeld and Andreas Pfitzmann, (2000) "Attacks on Steganographic Systems - Breaking the Steganographic Utilities EzStego, Jsteg, Steganos, and S Tools - and Some Lessons Learned," Lecture Notes in Computer Science, pp 61-75. https://pdfs.semanticscholar.org/4ffa/35072f8fab9efa7c17905c1735eeb502b1de.pdf?_ga=2.104924611.1927448487.1582298768-1200129864.1582211187