Steganography Using BCH-Codes

By

Muhammad Ahmad

Department of Mathematics
Quaid-i-Azam University
Islamabad, Pakistan
2018

Steganography Using BCH-Codes

By

Muhammad Ahmad

Supervised By

Dr. Asif Ali

Department of Mathematics
Quaid-i-Azam University
Islamabad, Pakistan
2018

Steganography Using BCH-Codes

By

Muhammad Ahmad

A Dissertation Submitted in the Partial fulfillment of the Requirement for the Degree of

MASTER OF PHILOSOPHY
IN
MATHEMATICS

Supervised By

Dr. Asif Ali

Department of Mathematics
Quaid-i-Azam University
Islamabad, Pakistan
2018

Preface

The improvement of the internet sources makes peoples bit by bit familiar with the computerized data through different systems. It is frequently evidence that an expensive piece of this information is either classified or private. Therefore, extraordinary security procedure has been used to give the required protection. So the issues of security and privacy have usually been approached using tools from cryptography. The message can be attached with a message authentication code (hash) and encrypted so that only the rightful recipient can read them and verify their integrity and authenticity. Modern cryptography is a developing field based on a demand mathematical foundation and decades of development.

Encrypted messages are obvious, and when a message is intercepted it is clear that the sender and recipient are communicating secretly. Steganography is the little and much younger sister of cryptography. It is an alternative tool for privacy and security. Instead of encrypting messages, we can hide them in other innocent looking objects so that their very presence is not exposed [24]. Thus, steganography can be a feasible alternative in countries where usage of encryption is illegal or in oppressive commands where using cryptography might attract unwanted attention. A recent example of a practical steganographic scheme that was used for information exchange between two subjects, one of which was residing at the time in a hostile country, was described by Toby Sharp at the 4th information hiding workshop [1].

Coding theory is a branch of mathematics which deals with the problems of errors that happens to the message when it is transferred through some communication channels. The goal of coding theory is to offer secure communication of information, in the sense that errors that occurred during the transmission can be corrected. However, to enjoy the benefit of this property some value must be paid, this value is in the form of redundant bits which are added in the transmitting data. There are two purposes of utilizing error correcting codes in steganography. The first purpose is to secure the secret information which is loaded in some other medium from different attacks such as reencoding and compression. The second purpose of utilizing error correcting code is to use the syndrome coding for data embedding. The purpose of using syndrome coding for data embedding is to find those bits in the cover medium such that there is a minimum distortion in cover medium. Here is some related work regarding steganography and coding theory.

In 2016 R J Mastafa utilize the concept of Discrete Cosine Transformation for video steganography along with BCH-Code [2]. For the increase in security of secret data, he encrypted the message first by using a key and then encoded with BCH (7,4,1) -Code. On the other hand, video is taken as a cover medium. Video is first converted into frames and then each frame is separated into Y, U, V components. The purpose of converting frames into Y, U, V space is to remove the correlation between colors, then the encoded message is embedded in these spaces. He uses BCH-Code prior to the embedding due to the motivation from the work of Liu et al [3]. Liu et al in his work discusses that (7,4,1) is the best code among the others.

In 2014 Ramadhan J Mstafa and Khalad M Elleithy presented a video steganographic algorithm they used uncompressed video sequences as a cover media. Before embedding the secret message is encoded first using Hamming (7,4) Code then this encoded message is embedded into the cover video [4].

In 2009, Eltahir et al presented a video steganography based on the Least Significant Bit (LSB). Author tried to increase the size of the secret message into the video frames. Video frames are used as a still image. A 3-3-2 approach has been used which means taking the LSB of all RGB color components (3 bits of red, 3 bits of Green, and 2 bits of Blue) [5].

So from the review of these and other related work we have proposed a new algorithm, “Video steganography using BCH-Code”. Chapter wise introduction of the thesis work is listed below.

In **chapter: 1** we have discussed some basic definitions of cryptography and steganography, which are useful for understanding of proposed work. This chapter is divided into three sections. Section one, covers some basic definitions regarding cryptography. Section two, covers basic definitions about steganography and in section three we have discussed classifications of steganography.

In **chapter: 2** we have discussed some basic definitions of coding theory and its applications in steganography. We have split this chapter into two sections. Section one, covers some basic definitions regarding coding theory and particular BCH-Code. In section two applications of steganography are discussed. At the end some related work is discussed

In **chapter: 3** we have discussed proposed work in detail In this article we have used Least Significant Bit (LSB) method for video steganography along with BCH-Code. Video is used as a cover medium and an encrypted image/text is used as a secret message. Firstly, encrypted image/text is converted into binary string and then encoded using BCH-Code. On the other hand, video is converted into frames and some of these frames are selected using a pseudo random number. At the end this encoded binary strings are embedded using LSB in each pixel of R, G, B plane of selected frames. It is observed that the proposed algorithm has high embedding efficiency and is robust against data loss. Chapter:4 includes the conclusion of the work.

Preface

The improvement of the internet sources makes peoples bit by bit familiar with the computerized data through different systems. It is frequently evidence that an expensive piece of this information is either classified or private. Therefore, the extraordinary security procedure has been used to give the required protection. So, the issues of security and privacy have usually been approached using tools from cryptography. The message can be attached with a "message authentication code" (hash) and encrypted. So that, only the rightful recipient can read them and verify their integrity and authenticity. Modern cryptography is a developing field based on a demand mathematical foundation and decades of development.

Encrypted messages are obvious and when a message is captured it is clear that the sender and recipient are communicating secretly. Steganography is the little and much younger sister of cryptography. It is an alternative tool for privacy and security. Instead of encrypting messages, we can hide them in other innocent looking objects so that their presence is not exposed [10]. Thus, steganography can be a feasible alternative in countries where usage of encryption is illegal or in oppressive commands where using cryptography might attract unwanted attention. A recent example of a practical steganographic scheme that was used for information exchange between two subjects, one of which was residing at the time in a hostile country, was described by Toby Sharp at the 4th information hiding workshop [21].

Coding theory is a branch of mathematics which deals with the problems of errors that happens to the message when it is transferred through some communication channels. The goal of coding theory is to offer secure communication of information, in the sense that errors that occurred during the transmission can be corrected. However, to enjoy the benefit of this property some value must be paid, this value is in the form of redundant bits which are added in the transmitting data [29]. There are two purposes of utilizing error correcting codes in steganography. The first purpose is to secure the secret information which is loaded in some other medium from different attacks, such as re-encoding and compression. The second purpose of utilizing error correcting code is to use the syndrome coding for data embedding. The purpose of using syndrome coding for data embedding is to find those bits in the cover medium so that there is a minimum distortion in the cover medium. Here is some related work regarding steganography and coding theory.

In 2016 R J Mastafa utilize the concept of Discrete Cosine Transformation for video steganography along with BCH-Code [19]. For the increase in security of secret data, he encrypted the message first by using a key and then encoded with BCH (7,4,1) -Code. On the other hand, a video is taken as a cover medium. The video is first converted into frames and then each frame is separated into Y, U, V components. The purpose of converting frames into Y, U, V space is to remove the correlation between colors. Then, the encoded message is embedded in these spaces. He uses BCH-Code prior to the embedding due to the motivation from the work of Liu et al [28]. Liu et al in his work discuss that (7,4,1) is the best code among the others.

In 2014 Ramadhan J Mstafa and Khalad M Elleithy presented a video steganographic algorithm. They used uncompressed video sequences as a cover media. Before embedding, the secret message is encoded first using Hamming (7,4) Code then, this encoded message is embedded into the cover video [18].

In 2009, Eltahir et al presented a video steganography based on the Least Significant Bit (LSB). Authors tried to increase the size of the secret message into the video frames. Video frames are used as a still image. A 3-3-2 approach has been used which means taking the LSB of all RGB color components (3 bits of red, 3 bits of Green, and 2 bits of Blue) [11].

So, from the review of these related works, we have proposed a new algorithm, “Video steganography using BCH-Code”. Chapter wise introduction of the thesis work is listed below.

In **chapter 1** we have discussed some basic definitions of cryptography and steganography, which are useful for the understanding of proposed work. This chapter consists of two portions. The first portion covers some basic definitions regarding cryptography. The second portion covers basic definitions about steganography

In **chapter 2** we have discussed some basic definitions of coding theory and its applications in steganography. We have split this chapter into three sections, Section one covers some basic definitions regarding coding theory and in section two we have discussed BCH-Code. In section three, applications of steganography are discussed.

In **chapter 3** we have discussed proposed work in detail. In this article, we have used the Least Significant Bit (LSB) method for video steganography along with BCH-Code. The video is used as a cover medium and an encrypted image/text is used as a secret message. Firstly, encrypted image/text is converted into a binary string and then encoded using BCH-Code. On the other hand, the video is converted into frames and some of these frames are selected using a pseudo-random number. In the end, these encoded binary strings are embedded using LSB in each pixel of R, G, B plane of selected frames. It is observed that the proposed algorithm has high embedding efficiency and is robust against data loss.

CONTENTS

Chapter 1

Preliminaries of Cryptography and Steganography

In this chapter, we have discussed some basic definitions of cryptography and steganography, which are useful for the understanding of proposed work. This chapter consists of two portions. First portion covers some basic definitions regarding cryptography. Second portion covers basic definitions about steganography[7,20,31].

1.1 Introduction to Cryptography

1.1.1 Cryptology

Cryptology is imported from the Greek word “Kryptos” which means covered up. It is related to the study of the modern cryptosystem. Cryptology is divided into two main branches cryptography and cryptanalysis.

1.1.2 Cryptography

Cryptography is the art and science of writing secret messages into an incomprehensible form so that unauthorized individuals would not understand the contents of the secret message. It is the study of making secure cryptosystem.

1.1.3.Cryptanalysis

Cryptanalysis is the branch of Cryptology which deals with the study of breaking cryptosystem. It is an important part of cryptology. Because to check the reliability of cryptosystem cryptanalysis is very essential.

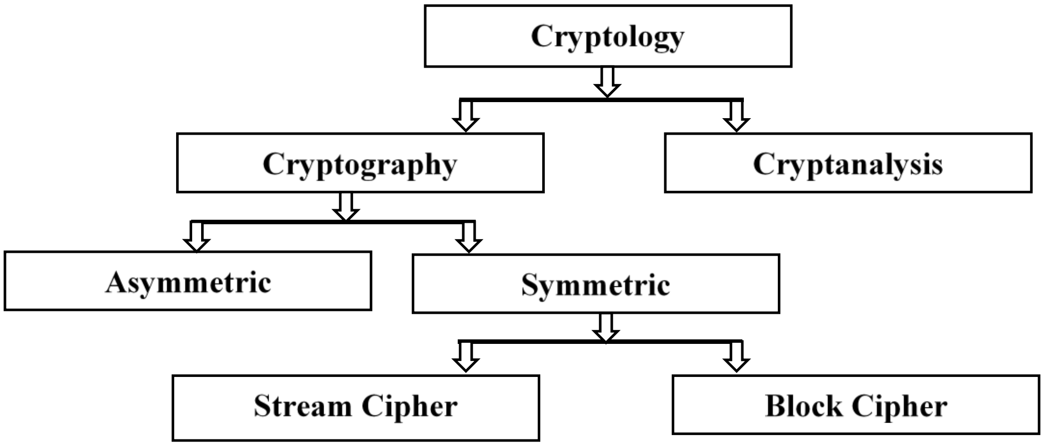


Figure 1 types of cryptography

1.2 Types of Cryptography

Modern cryptography consists of three categories depending on the key used for encryption and decryption.

1) Symmetric key cryptography.
2) Asymmetric key cryptography.

1.2.1 Symmetric Key Cryptography

Symmetric key cryptography consists of the following map

$$E : K \times P \to C$$

Such that the map $E_k : P \to C \ , p \mapsto E(k, p)$ is invertible for each $k \in K$.

Where E_k is the encryption function corresponding to the key k and P is the set of plain text C is the set of cipher text.

In symmetric key cryptography, both parties use the same key for encryption and decryption as shown in fig. Both the parties kept the key secret and share with each other using a secure channel. Encryption algorithm E and decryption algorithm D are made public. The main problem in symmetric key cryptography is to share a secret key. Because if anybody gets the key he can easily decrypt the message. Due to the key difference, symmetric key cryptography is further divided into two branches.

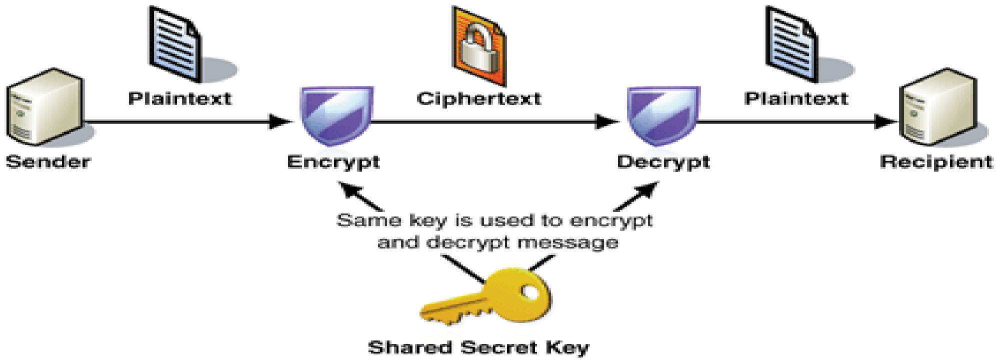


Figure 2 Symmetric Key Cryptosystem.

Stream Ciphers: [31] Consider a set K of keys and a set P of plain text. Then a stream cipher can be described as follows.

$$E^{\circ}: K^{\circ} \times P^{\circ} \to C^{\circ}$$

This map encodes a stream $p \coloneqq p_1 p_2 p_3 \cdots \in P^{\circ}$ into a stream $c \coloneqq c_1 c_2 c_3 \cdots \in C^{\circ}$ by utilizing a stream of keys $k \coloneqq k_1 k_2 k_3 \cdots \in K^{\circ}$.

Block Ciphers: [31] It is another type of symmetric key cryptographic scheme with $P = C = \{1,0\}^r$ and $K = \{1,0\}^s$ Such that

$$E: \{1,0\}^s \times \{1,0\}^r \to \{1,0\}^r$$

Where E is the encryption algorithm which encrypts a plain text of binary length r in a cipher text of binary length r by using a key having binary length s. Examples of block ciphers are AES and DES.

1.2.2 Asymmetric Key Cryptography

Symmetric key cryptography offers security in communication between the two parties. But communication between two parties remains secure until the secret key is switched using a secure channel. So, the public key cryptography provides security for key exchange. In public key cryptography, two different keys are used for encryption and decryption. One of these keys is secret key and other is the public key. The first public key cryptographic algorithm was designed by Shamir, Rivest, and Adleman in 1978 and is named as RSA. RSA is based on the factorization of large primes.

RSA Encryption: [31] If a plain text p and public key $k = (n, e)$ Is given, then the encryption map can be defined as follows.

$$q = e_k(p) = p^e \: mod \: n$$

Where $p, q \in Z_n$.

RSA Decryption: [31] If a cipher text q and private key $k = d$ are given, then the decryption map can be defined as

$p = d_k(q) = q^d mod \: n$ Where $p, q \in Z_n$.

1.3 Basic Terms of Cryptography

Some basic terminologies used for cryptography are as follows [20].

1.3.1 Plain Text

A text which is in a comprehensible form and can be understood without any extra information is called plain text.

1.3.2 Cipher Text

A text which is in incomprehensible form and can't be read without extra information is called cipher text.

1.3.3 Ciphers

It is an algorithm which is utilized to convert plain text into cipher text.

1.3.4 Encryption

It is the scheme to transform a plain text into cipher text. The mathematical expression of this transformation is.

$$c = e(p)$$

Where c is ciphered text e is encryption function and p is plain text.

1.3.5 Decryption

It is the scheme to renovate cipher text into plain text. Mathematically, this transformation is represented as.

$$p = d(c)$$

Where p denotes plain text, c denotes cipher text and d denotes decryption function.

1.3.6 Key

Some information is required to transform plain text into a cipher text and vice versa. This information is called a key.

Remarks: If we have a cryptosystem consisting of five parameters (K, M, C, E, D) Where K denotes the key, M denotes the plain text, C denotes the cipher text, E denotes the encryption function and D denotes the decryption function. Then corresponding to each $k \in K$ an encryption function e_k and decryption function d_k exist such that $d_k\left(e_k(x)\right) = x$. Where $x \in M$.

1.4 Security Services of Cryptography

Security services are the facilities provided by cryptographic algorithms for secure communication. These services ensure secure communication through insecure channels. Some of these security services are given [20,31].

1.4.1 Authentication

This service ensures the authenticity in communication. It conforms to the receiver that the message he has received is sent by a documented and confirmed sender [31].

1.4.2 Data Integrity

This service ensures that the message is received at the receiving end without any distortion and modification. For this purpose usually, a hash function is used [31].

1.4.3 Confidentiality

This service makes sure that the content of secret information can’t be read by an unauthorized party. To achieve this service secret information is encrypted using symmetric key cryptography.

1.4.4 Availability

Information will be useless if an authorized entity can’t get access and control to his data. Availability assures that when an authorized person demands he can access to his personal data.

1.5 Introduction to Steganography

The term steganography is derived from Greek words “Steganos” which means covered and “graphy” which means writing thus steganography means covered writing. steganography is an art and science of secret communication. The first confirmation about the steganography is being used for secret communication is written by Herodotus. He tells the story of slaves guided by their chief Histiaeus with a secret message tattooed on the scalp of the slaves. After some time, slave’s hair grew up and the secret message is concealed. On reaching their destination their head is shaved again to reveal the secret message. He also documented the story of Demeratus who inform Sparta about the plan of Persian great king Xerxes. Demeratus scratched his caution onto the tablet and the tablet was covered again with a fresh coat of wax so that tablets seem to be blank. These blank tablets are carried to Sparta without any suspicion [7].

Many steganographic techniques were used at that time such as message carried by pigeons or hidden on women’s earring [26]. At that time Acrostic was one of the most popular methods for steganography. In this method letters of secret message do not form systematic structure and appear as a random shape [27]. A secret message hidden using this method can be read simply by placing a cover over the text. This cover plays a role of stego key which is shared among the communicating parties. Acrostic steganographic method was used by Germans and Allies in World War 1. Modern steganographic techniques were introduced by Brassil [4]. According to him shifting of lines up or down by $\frac{1}{300}$ in an inch is not visually noticeable. Another idea which was proposed by Brewster (1857) played a central role in several wars in the nineteenth and twentieth centuries. He suggested that hidden message may shrink enough so that it resembles with space of dirt. The shrinking was done by using a technology introduced by French photographer Dragon in the Franco Perussion war. In world war 1 the Germans used a microscopic image (microdots) and hide them in the bends of postcards. The use of microdots was discovered by the Allies in 1941.

1.5.1 Importance of Steganography

Electronic communication is gradually prone to snooping. The security issues can be approached using cryptographic tools. With the help of these tools, the message can be encrypted using an authentication code so that only correct person can read and confirm integrity and authenticity. Modern cryptography is an advanced field based on mathematical fundamentals. Cryptography is closely linked with steganography. Steganography is a substitute for security and privacy. Using steganography, we can hide a message in other objects so that the presence of the message would not be exposed. Thus, steganography is a better choice in the countries where encryption is considered illegal [7].

1.5.2 Applications of Steganography

The reasons for secret communication between two parties are many, for example, if two lovers want a secret relationship or in the case of organizations who wants to communicate with other organizations outside the country. Criminals also use steganography to organized crimes. So, there are lots of applications of steganography. An interesting use of steganographic algorithm was applied by the United States and the Soviet Union, both the countries placed sensors in nuclear services to communicate the number the number of missiles without sharing any other information [22]. To avoid unauthorized access from the transmitted signals, communication was made secure using steganographic. Here we discuss two important applications of steganography, which are secret communication by dissidents and concealed communication by criminals.

Steganography for Insurgents

Many countries in the world have rebels which are not legal. Any such type of organization takes care when communicating with each other or with other organizations outside the country. As such rebels are fully aware that their activities may be under the eyes surveillance and opponents. So, the only way that their communication might be secret is the use of cryptography or steganography. Using cryptography dissidents can communicate with each other securely. But the drawback is that if two people exchanging encrypted message are caught by adversary then they will be arrested along with the encrypted message and the secret key. Steganography is the only way for secure communication between dissidents. Science basic communication is essential for the modern country. So, every country has permeated to communicate, thus dissidents can also communicate securely[7].

Steganography for Criminals

In a democratic country, every citizen is monitored by the government. For example, in the United States, phone calls of every person is monitored with the permeation of a federal judge's order. In this environment, an individual would not aware that he is under surveillance. So, the only way to communicate securely with each other is to use steganography.

1.6 Properties of Steganography

It is very difficult for a steganographic scheme that it covers all the properties which are required for good steganography. Here we discuss some of the properties [7].

1.6.1 Capacity

It is defined as the maximum amount of data that can be hidden inside a cover medium.

1.6.2 Efficiency

It is defined as the number of secret bits that can be embedded in a cover medium with the per unit of alteration.

Remarks: Embedding efficiency and embedding payload are inverses of each other. If we want to increase embedding payload we must decrease embedding efficiency. The security of any steganographic algorithm directly depends upon embedding efficiency.

1.6.3 Security

Any steganographic algorithm will be secure. If the embedded data cannot be ejected by the eavesdropper after recognition. It depends upon the information about the key and embedding algorithm.

1.7 Components of Steganography

Some basic components for steganography are.

1.7.1 Cover Medium

The digital data in which we hide our secret message is called cover medium. Digital data may be in the form of Audio, Text, Image, Video.

1.7.2 Secret Message

Any type of digital secret information which we embed inside a cover medium is called a secret message.

1.7.3 Stego Medium

Cover medium containing the secret message is called a stego medium.

1.7.4 Key

Some extra information needed for embedding and extracting purpose is called a key.

1.7.5 Embedding Algorithm

An algorithm which is required for embedding secret information inside a cover medium.

1.7.6 Extracting Algorithm

The inverse process for extracting secret information from the cover medium.

1.8 Types of Steganography

1.8.1 Pure Steganography

The steganographic scheme in which we do not have need key for embedding and extracting purpose is called pure steganography. In pure steganography, four parameters are required $(C_{cov}, S_{sec}, E_{emb}, D_{ext})$ Such that $|C_{cov}| \geq |S_{sec}|$ E_{emb}: $C_{cov} \times S_{sec} \rightarrow C_{steg}$ and $D_{ext}: C_{steg} \rightarrow S_{sec}$.

1.8.2 Symmetric Key Steganography

The steganographic scheme in which we use the same key for embedding and extracting process is called symmetric key steganography.

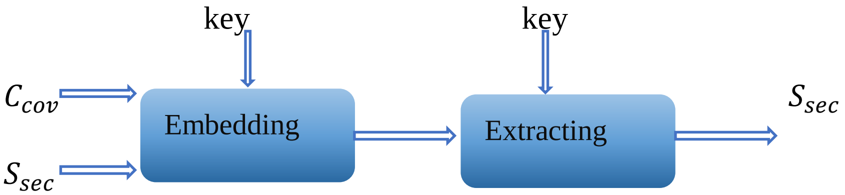


Figure 3 Symmetric key Steganography

1.8.3 Asymmetric Key Steganography

The steganographic scheme in which two different keys are required for embedding and extracting secret information is called asymmetric key steganography. One is the public key and other is the private key.

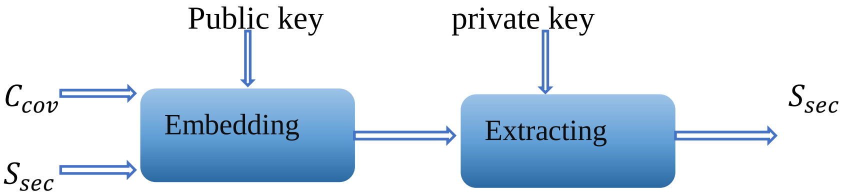


Figure 4 Asymmetric key Steganography

1.9 Classification of Steganography

There are many kinds of digital files which can be used as a cover medium. But the most suitable file formats are those which has a large number of redundant bits. Redundant bits are those bits which can be altered without any noticeable change in the cover file. Among these files formats, Images and Videos have a large degree of redundant bits. Steganography is classified according to the cover medium used for steganography.

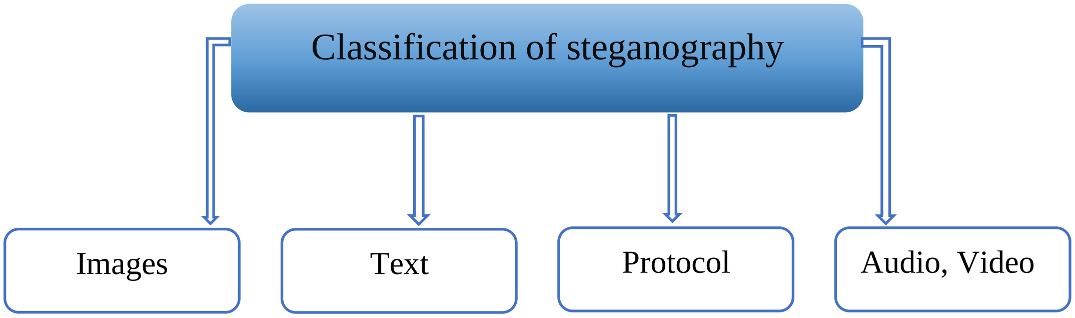


Figure 5 Classification of Steganography

1.9.1 Image Steganography

In image steganography, we use digital images as a cover medium.

Digital image

Digital images are made up of very small components called pixel or picture element. These pixels contain some information about the intensity of colors in the form of numbers. In a computer, these numbers are represented in the form of bits 0 mean low or black and 1 means high or white. A binary image pixel has two values only either 0 or 1. In 8-bit grayscale image intensity of gray color lies between 0 and 255. In the 24-bit color image, each pixel consists of 3-byte 1 byte represents 1 basic color so in a single pixel there is 3 basic color named as Red, Green, Blue.

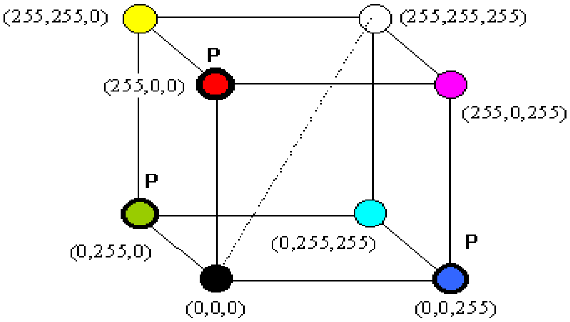


Figure 6 Color model of a digital image

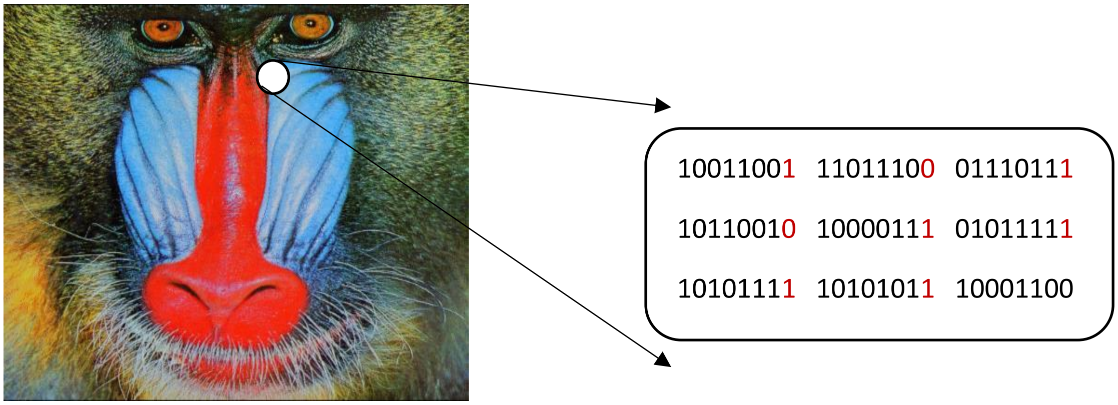
Images are the most favorite medium for steganography. Because images consist of many redundant bits. There are two main techniques used for image steganography. Spatial domain technique and transformation domain technique.

Steganography Using the Spatial Domain

Here are some schemes used for image steganography in the spatial domain. (1) Steganography using LSB (2) Steganography using the difference between pixel values (3) College based steganography.

1) Steganography using LSB

The LSB is a simple approach for hiding secret information in images. In this technique, the 8th bit of each byte is replaced with the most significant bit of the secret message. If we have a desire to increase the capacity of embedding data, then we must replace lest 2 significant bits of the cover image with the bits of the secret message. In the same way, payload can be increased by replacing 3 or 4 least significant bits. But then due to the increase of data for embedding quality of the cover image will be low. For example, if we want to embed 1 byte=10101111 of the secret image in the cover image of Baboon then we must select 3 pixels of the Baboon image as follows [1,15,2].



10011001	11011100	01110111
10110010	10000111	01011111
10101111	10101011	10001100

Figure 7 Least Significant Bit

Steganography using Frequency Domain

In frequency domain pixel value of an image is transformed from spatial domain to the frequency domain usually, three types of transformations are used for this purpose (1) Discrete Cosine Transformation (2) Discrete Fourier Transformation (3) Discrete Wavelet Transformation.

1) Discrete Cosine Transformation (DCT)

DCT is being used extensively for watermarking and steganography. It is widely used for image processing such as video and image compression [2,12,24].

Discrete cosine transformation converts the pixel values of an image in high, middle, and low-frequency components. Due to the less importance of high-frequency signal component, it is usually removed. Maximum signal energy lies in the low-frequency component it is the most visual part of the image so secret data is not embedded in this part. Middle frequency components are useful for embedding therefore much of the data is embedded in this part. Consider a $M \times N$ image to transform images from spatial to the frequency domain and reverse from frequency to spatial domain following equations are used respectively.

$$F_{rs} = \alpha_r \alpha_s \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} A_{pq} \; cos\left(\frac{\pi(2p+1)r}{2M}\right) cos\left(\frac{\pi(2q+1)s}{2N}\right)$$

$$S_{pq} = \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} \alpha_r \, \alpha_s F_{rs} \, cos\left(\frac{\pi(2p+1)r}{2M}\right) \; cos\left(\frac{\pi(2q+1)s}{2N}\right)$$

Where $\alpha_r = \begin{cases} \frac{1}{\sqrt{M}}, & r = 0 \\ \sqrt{\frac{2}{M}}, & 1 \leq r \leq M-1 \end{cases}$

And $\alpha_s = \begin{cases} \frac{1}{\sqrt{N}} & , \quad s = 0 \\ \sqrt{\frac{2}{N}} & ,1 \leq s \leq N-1 \end{cases}$

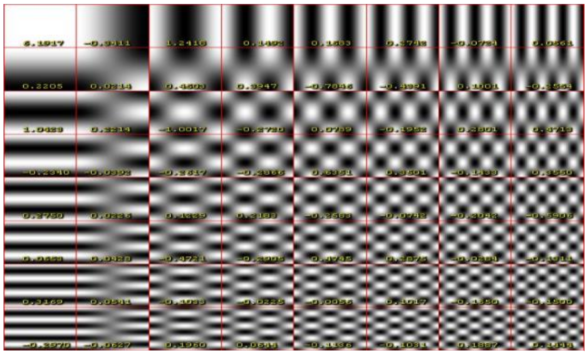


Figure 8 Discrete Cosine Transformation.

1.9.2 Audio Steganography

In audio steganography secret data are embedded in the audio files and the audio file may be of the following type Mp3, WAV,[6,8] Au. Many techniques exist for audio steganography some of these techniques are as follows. (1) Echo Hiding (2) Phase Coding (3) Parity Coding (4) Spread Spectrum (5) Tone Insertion (6) Least Significant Bit.

1) Echo Hiding

This method of hiding is used to hide a secret message in an audio by passing echo from a discrete signal of the frame. This technique is robust as compared to other methods, but the drawback is that this technique has a low capacity of hiding secret data.

2) Phase Coding

In this technique selected components of the phase from the cover audio signal spectrum are replaced with secret information. Phase components should be kept small so that the new amplitude is not noticeable by the human.

3) Spread Spectrum

In this technique encoded secret information is spread out into the entire available frequency signals.

4) Least Significant Bit

In this technique, LSB of each audio frame is replaced with the bits of secret information. It is the simplest method with high embedding payload.

1.9.3 Text Steganography

A need of protecting important information is a major issue for intelligence agencies and companies for their business. For the security of important information, a large number of techniques are available depending upon the cover media used for this purpose. Unlimited digital files are available for the use as a cover media. The text is the oldest cover media used for steganography. Text as a cover media is not commonly used due to the less number of redundant bits. Here we discuss some techniques used for text steganography [14,16].

1) Word Spelling

This technique of steganography is used for English text only because word spelling used in the US is different from the spelling used in the UK. For example, in US favorite and UK favourite there is a difference of u.

2) Semantic Method

In this method synonym of the words are used for steganography. This method is better than the previous one.

3) Abbreviation

This technique is used to hide information in the text by changing words with their abbreviations. Using this method, a very little amount of secret information can be embedded.

4) Line Shifting

In this technique lines of cover, the text is shifted vertically to some degree, so that the information is embedded by forming a unique form of a text.

1.9.4 Protocol Steganography

There are many cover mediums used for the steganographic purpose. Protocol steganography is a modern technique in this technique network protocols are used for the steganographic purpose [9]. To hide secret information reserved and redundant bits of the protocols and header fields are used for embedding. Here are some methods for protocol steganography. (1) Padding steganography (2) transcoding steganography (3) torrent steganography (4) Skype hides (5) suggested steganography (6) lost audio packet steganography. (LACK). We discuss here some of these techniques.

1) LACK

In this technique, voice packets are produced on the transmitter side. Out of these packets, one is selected, and the payload is replaced with the secret information. As the delay timer detects the packet, it is sent to the receiver side.

2) Suggested Steganography

When we search for something online with google. Google gives us some suggestions by itself secret information is concealed by injecting a letter attached with each suggested word.

3) Skype Hides

Skype is a P2P transmission service. In this service, there are many packets having no voice signals. In this technique, these packets are replaced with the secret information with the minimum possible distortion.

1.10 Image Compression

When we are working with the images of larger bit depth, the size of images becomes enough large, so that transmission of images over the internet becomes too difficult. Images of larger size take much time to display. To overcome this difficulty some techniques are required to reduce the size of images. These techniques use some mathematical formulas such as discrete cosine transformation. Images are compressed using two types of techniques. Lossy compression and lossless compression techniques.

1.10.1 Lossy Compression

The purpose of both methods of compression is to save the storage space, but both techniques have different procedures and importance. In lossy compression, less significant data from the original image is discarded so that the size of the new file is small as compared to the original file. Due to the lossy compression, the visual quality of the compressed image will be low as compared to the original image.

1.10.2 Lossless Compression

In lossless compression, technique data is converted into some mathematical form and is not removed from the original image. So that the visual quality of the decompressed image and the original image will be same.

1.11 Analysis for Steganography

Here we discuss some analysis for steganography.

1.11.1 Mean Square Error (MSE)

The mean square error is used to check the consistency of the algorithm. It is calculated between stego and original image with the help of the following equation.

$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (O(i,j) - S(i,j))^2$$

Where $M \times N$ represents the size of image and $O(i,j), S(i,j)$ represents the pixel location at ith row and jth column of original and stego image respectively.

1.11.2 Root Mean Square Error (RMSE)

RMSE is used to calculate the accuracy and robustness of the stego image. The error between the original and stego image is calculated by using the following equation. It is the square root of the mean square error.

$$RMSE = \sqrt{\frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (O(i,j) - S(i,j))^2}$$

1.11.3 Peak Signal to Noise Ratio (PSNR)

To calculate the visual quality of the image PSNR is calculated. With the help of the following equation, we can calculate the PSNR of the image.

$$PSNR = 10log_2 \left(\frac{f_{max}^2}{MSE} \right)$$

Where f_{max} is the maximum pixel value of an image and MSE is the mean square error of the image. PSNR value is measured in dB if PSNR value is greater then 30dB then it will be difficult for the intruder to find the difference between original and stego image.

Chapter 2

A Class of Cyclic Codes and its Applications in Steganography

In this chapter, we have discussed some basic definitions of coding theory and its applications in steganography. We have split this chapter into three sections. Section one covers some basic definitions regarding coding theory. In section two BCH-Codes is discussed and in section three applications of steganography are discussed [29,30].

2.1 Introduction to Coding Theory

Coding theory is a branch of mathematics which deals with the problems of errors that occurs in the message when it is transmitted through some communication channels. These channels might be television, telephone, internet, radio, and recording devices. The error in the message may be due to faulty apparatus, thermal noise or human negligence. A channel which produces an error during transmission is called a noise channel. Error correcting code is a scheme which encodes the message before transmission from the noisy channel and if an error occurs in the message during transmission, then the decoding scheme will correct the error. In general, error correcting code adds redundant bits in the message to find and correct the error bit in the message.

2.1.1 Code

Consider a set of finite elements q is denoted by S and the set consisting of all m-tuples of elements of S are denoted by $\boldsymbol{S}^m$ where $m>1$ and belongs to a positive integer. Then there are q^m elements in $\boldsymbol{S}^m$ these elements are called words or vectors. If $C\neq\Phi$ is a subset of $\boldsymbol{S}^m$, then C is named as q-ary code of length m over $\boldsymbol{S}$. In general, if the value q=2 then C is named as binary code and if q=3 then C is named as ternary code. The entries of C are called codeword. If entry of C is only one, then C is called trivial codeword and if entries of C are of the form $vvvv...$v. where v belongs to some $\boldsymbol{S}$ then C is named as repetition code.

Example 1

If C= {000 111} then C is named as binary repetition code of length 3.

2.1.2 Hamming Distance

Let *u*,*v* ∈ $\boldsymbol{W}^n$ where *u*=$u_1u_2u_{3\ldots}u_n$,$v = v_1v_2v_3 \ldots v_n$. Then the Hamming Distance between *u* and *v is* denoted by $d = (u, v)$ And is defined as.

$$d(u, v) = |\{i|u_i \neq v_i\}|$$

Example 2

Let $u = 11011$ and$v = 10101$ then the Hamming Distance is 3 i.e. $d(11011{,}10101) = 3$.

2.1.3 Minimum Distance

Minimum distance between two different codewords in *C* is denoted by $d(C)$ and is defined as

$d(C) = min\{d(u, v)|\ u, v \in C, u \neq v\}$.

Theorem 1

Suppose a code *C* has a minimum distance *d* and let $t = \left\lfloor\frac{d-1}{2}\right\rfloor$ Then

(1) *C* has the capability to detect $d - 1$ error in transmitting codeword.

(2) *C* has the capability to correct at the most *t* error in transmitting codeword.

2.1.4 Perfect code

A code *C* is said to be a perfect code if $C \subset U^n$ and having least distance $d(C) = 2t + 1$. If $\forall\ u \in U^n$ There exist $v \in C$such that $d(u, v) \leq t$.

Example 3

Consider a code $C = \{0000{,}1111\}$ This is a perfect binary code of length 3.

2.1.5 Linear Code

Suppose we have a finite field *F* and $n \in Z^+$, And if $C \subset F^n$. Then *C* is said to be a linear code. If the dimension of *C* is *k*, then *C* is said to be $[n, k]$ Code. And if *d* is the minimum distance, then *C* is called an $[n, k, d]$ Code.

Example 4

Consider $C = \{000{,}111\}$ Then *C* is a [3,1,3] linear repetition code over F_2.

2.1.6 Generator Matrices

Consider a linear code *C* having length *n* and dimension *k* and let a matrix *G* of order $k \times n$ such that rows of *G* form bases of the code *C* then the matrix *G* is called the generator matrix of *C*.

2.1.7 Dual Code

Consider a code *C* of length *n* and dimension *k* over a finite field *F*. Then we define a dual code of *C* as.

$$C^{\perp} = \{v \in F^n |\, u \cdot v = 0 \;\forall\, u \in C\}$$

2.1.8 Equivalent Codes

Suppose C, C' are linear codes of length *n* and dimension *k* over a field *F*. *If* ∃ a mapping $\varphi: C \to C'$ which is bijective then these codes are called equivalent codes.

$$\varphi\big(y_{1,}y_{2,} \cdots, y_n\big) = (\beta_1 y_{\sigma(1),} \cdots, \beta_n y_{\sigma(n)})$$

Where $\beta_{1,} \cdots, \beta_n \in F$ are scalars and σ is a permutation of $\{1,2, \cdots, n\}$.

2.1.9 Hamming Codes

Consider $1 \neq s \in Z^+$ and let a matrix $[H]_{s\times(2^s-1)}$ columns of this matrix are distinct and non-zero vectors from F_2^s. Then code having parity check matrix *H* are called binary hamming code which is denoted by $Ham(s, 2)$. So, for a fixed *s* there will be $(2^s - 1)!$ binary hamming codes which are equivalent. As *H* is an $s \times (2^s - 1)$ matrix. Hence $Ham(s, 2)$ is a code having length $n = 2^s - 1$ and dimension is $k = n - s = 2^s - 1 - s$ here value of $s = n - k$ tells the information about redundant bits in the code.

Example 5

If we take, $s = 3$ then $Ham(3,2)$ is simply a [7,4] Code having parity check matrix *H*. Length of this code is $n = 2^3 - 1 = 7$ and dimension is $k = 2^3 - 1 - 3 = 4$ and the codewords are$2^4 = 16$.

2.1.10 Cyclic Shift

The linear mapping $\rho: F^n \to F^n$ which is given by.

$$\rho(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_n, \alpha_1, \cdots, \alpha_{n-1})$$

Is known as cyclic shift.

2.1.11 Cyclic Codes

Suppose a linear code *C* which is a subspace of F^n Then the code *C* is known as cyclic code if

$$\rho(\alpha) \in C , \forall \alpha \in C$$

Remarks: Consider a set containing all polynomials in *x* having degree smaller than *n* over the field *F* is denoted by.

$$F[x]_n = \{b_0 + b_1x + \cdots + b_{n-1}x^{n-1} | b_{0,}b_1, \cdots , b_{n-1} \in F\}$$

This set of polynomials is a vector space having dimension *n* over the field *F*. Since the mapping from the vector space F^n to the vector space $F[x]_n$ Is isomorphic, so we can treat every vector in F^n as a polynomial in $F[x]_n$. If we have a polynomial ring, $F[x]$ then this polynomial ring becomes a field iff when we take the quotient of this polynomial ring with the irreducible polynomial $P(x) \in F[x]$ Further, if the degree of the irreducible polynomial is *n* then

$$\frac{F[x]}{\langle p(x)\rangle} = \{b_0 + b_1t + \cdots + b_{n-1}t^{n-1}|b_0, b_1, b_2, \cdots , b_{n-1} \in F\}$$

Where t is the coset $x + \langle p(x)\rangle$ So that $P(t) = 0$.

In the case when $P(x) = x^n - 1$ then the quotient ring is not a field. The reason is that $P(x) = x^n - 1$ this is not irreducible.

2.1.12 Generator Polynomial of the Cyclic Code

Let $C \neq 0$ be an ideal in $F[x]_n$. Suppose $g(x)$ be a unique and monic polynomial with the smallest degree in *C*. Then $g(x)$ is called a generator polynomial of *C*.

Theorem 2

Consider a cyclic code *C* which is a subspace of $F[x]_n$ having generator polynomial

$$G(x) = a_0 + a_1x + \cdots + a_sx^s$$

Where $a_s = 1$ then the dimension of *C* is $n - s$.and the generator matrix of *C* is

$$G = \begin{bmatrix} a_0 & a_1 & \cdots & a_s & \cdots & 0 \\ 0 & a_0 & \cdots & a_{s-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_0 & \cdots & a_s \end{bmatrix}$$

The order of this matrix is $(n-s) \times n$.

2.1.13 Check Polynomial

Consider a cyclic code $C \subset F[x]_n$ with generator polynomial $g(x)$. Then a polynomial $h(x)$ such that $x^n - 1 = g(x)h(x)$ is said to be check polynomial of the code *C*.

2.2 BCH-Codes

A very superior kind of cyclic codes is BCH-Codes. Before discussing BCH-Codes, we must understand some properties of irreducible polynomial and finite fields. Since corresponding to every prime and some positive integer *s* there is a single field of order p^s and is denoted as F_{p^s}. If $s = 1$ then we have $F_p = Z_p$. Suppose $F = F_q$ where $q = p^s$ Then set of all nonzero elements of *F* form cyclic group under multiplication, the order of this set is $q-1$ and denoted by F^*. Hence, for every $x \in F^*$ we have $x^{q-1} = 1$ that is $x^q = x$. Now corresponding to any divisor *m* of $q-1$ we have an element $x \in F^*$ such that $o(x) = m$. this element *x* is said to be the primitive nth root of unity in *F*. Further, if the order of *x* is $q-1$ then *x* is a generator of the cyclic group F^* this *x* is said to be a primitive element of *F*. Science characteristic of *F* is p.so $\forall\ x \in F\ px = 0$ and also $qx = 0$.

If we have a finite field F_q. Then we can construct a new field F_{q^r} from the quotient ring $\frac{F_q[x]}{\langle p(x) \rangle}$ where $p(x)$ is the irreducible polynomial with degree r in $F_q[x]$. This field F_{q^r} is said to be extension of F_q [30].

2.2.1 Minimal Polynomial

Let $\beta \in F_{q^r}$ then there is a monic irreducible polynomial of minimum degree $g(x) \in F_q[x]$ such that $g(\beta) = 0$ then this polynomial is said to be a minimum polynomial of β over F_q.

Theorem 3

Let $\beta \in F_{q^r}$ then $\beta, \beta^q, \beta^{q^2}, \cdots$have identical minimal polynomial over the field F_q.

Now we define the BCH-Code as follows. Let $n, c, q, d \in Z^{+}$ such that$n \geq d \geq 2$, q is the power of some prime and q, n is relatively prime to each other. Consider a smallest positive integer *r* be such that $q^{r} \equiv 1(mod\, n)$. Then by using Euler, s theorem we have n divides $q^{r}-1$ if the primitive nth root of unity is η in $F_{q^{r}}$. Let $h_{j}(x) \in F_{q}[x]$ is the minimal polynomial corresponding to η^{j}. Let us denote the product of all different minimal polynomials $h_{j}(x)$, where $j = c, c+1, c+2, \cdots c+d-1, c+d-2$; by $g(x)$. Since $h_{j}(x)$ divides $x^{n}-1$ for each *j* then $g(x)$ also divides $x^{n}-1$.

Now if the cyclic code *C* generated by the polynomial $g(x)$ then the cyclic code *C* is known as BCH code having length *n* and designed distance *d*.

Note that if $n = q^{r}-1$ then the code *C* is known as primitive BCH-code and if *c=1* then *C* is said to be a narrow sense BCH-Code.

Theorem 4

Suppose a BCH code *C* of length *n* and designed distance *d* over F_{q}. Then

$$C = \left\{ v(x) \in F_{q}[x]_{n} \mid v\left(\eta^{j}\right) = 0 \ \forall \, j = c, c+1, \cdots, c+d-2 \right\}$$

Equivalently the code *C* is a null space of the following matrix.

$$H = \begin{bmatrix} 1 & \eta^{c} & \eta^{2c} & \cdots & \eta^{(n-1)c} \\ 1 & \eta^{c+1} & \eta^{2(c+1)} & \cdots & \eta^{(n-1)(c+1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \eta^{c+d-2} & \eta^{2(c+d-2)} & \cdots & \eta^{(n-1)(c+d-2)} \end{bmatrix}$$

Theorem 5

Consider a BCH-code *C* with designed distance. d Then we have $\mathrm{d}(C) \geq \mathrm{d}$.

2.2.2 Encoding of BCH-Code

To transmit a message through any channel. The message is first converted into blocks of *k* bits to encode a message $n-k$ bits are added in each block to form a codeword these extra bits are called redundant bits. Then this codeword is transmitted through noisy channels. If from the receiver side, we get the same codeword then no error occurs during transmission. But if we cannot get the codeword from the receiver side that's mean an error has occurred during transmission, we encode the message by using the following steps.

(1) First, a k bit message block is converted into a polynomial.
(2) Then we find the generator polynomial.
(3) To get a code word we multiply message polynomial with a generator polynomial.

Here we explain the method of finding generator polynomial with an example.

Example 6

Construct a narrow sense binary BCH-code with length n=15 having designed distance d=7.

Since we have a binary BCH-code so $q = 2$ and $m = 4$ because of $n = 15 = 2^4 - 1$ we take a polynomial $p(x) = x^4 + x + 1$ Which is primitive, irreducible over F_2 then the field F_{2^4} is represented by

$$F_{2^4} = \{b_0 + b_1 + b_2\eta^2 + b_3\eta^3 \,|b_0, b_1, b_2, b_3 \in F_2 \}$$

Where η satisfies the equation $\eta^4 + \eta + 1 = 0$ using this equation, we can find all powers of η as follows.

$\eta^4 = 1 + \eta$	$\eta^8 = 1 + \eta^2$	$\eta^{12} = 1 + \eta + \eta^2 + \eta^3$
$\eta^5 = \eta + \eta^2$	$\eta^9 = \eta + \eta^3$	$\eta^{13} = 1 + \eta^2 + \eta^3$
$\eta^6 = \eta^2 + \eta^3$	$\eta^{10} = 1 + \eta + \eta^2$	$\eta^{14} = 1 + \eta^3$
$\eta^7 = 1 + \eta + \eta^3$	$\eta^{11} = \eta + \eta^2 + \eta^3$	$\eta^{15} = 1$

Now to find a BCH-Code with designed distance 7 we have need minimal polynomial corresponding to η^j. Where $j = 1,2,\cdots,6$ from the theorem η, η^2, η^4 have the identical minimal polynomials. If we let $h(x)$ is the minimal polynomial corresponding to η^3 then $\eta^3, \eta^6, \eta^{12}, \cdots$also have $h(x)$ as the minimal polynomial. From the relation $\eta^{15} = 1$ we note that $\eta^3, \eta^6, \eta^9, \eta^{12}$ are the roots of $h(x)$.

$$h(x) = (x - \eta^3)(x - \eta^6)(x - \eta^9)(x - \eta^{12})$$

$$= x^4 - (\eta^3 + \eta^6 + \eta^9 + \eta^{12})x^3 + (\eta^3 + \eta^6 + \eta^9 + \eta^{12})x^2 - (\eta^3 + \eta^6 + \eta^9 + \eta^{12})x +$$

$$= x^4 + x^3 + x^2 + x + 1$$

Similarly, $q(x)$ is the minimal polynomial corresponding to η^5 having roots η^5, η^{10}

$$q(x) = x^2 + x + 1$$

Thus, generator polynomial is

$$g(x) = x^{10} + x^8 + x^5 + x^4 + x^2 + x + 1$$

2.2.3 Decoding of BCH-Code

We use the following steps to decode the received message.

(1) First, we will find the syndromes by using the received polynomial.

(2) Then calculate $\sigma(x)$ Where $\sigma(x)$ is the error locator polynomial.

(3) By calculating the inverse of the roots of $\sigma(x)$ we get the error location number.

Suppose a codeword $v(x)$ is generated from the encoder side and it is transmitted through the noisy channel. If an error occurs during transmission, then codeword $v(x)$ Converts into invalid codeword $r(x)$. Suppose $e(x)$ is the error pattern then.

$$r(x) = v(x) + e(x)$$

By using the relation, we can calculate the syndrome.$S_j = r(\eta^j) = v(\eta^j) + e(\eta^j)$

Because of $v(x)$ is a valid codeword so $v(\eta^j) = 0$. Therefore $S_j = e(\eta^j)$. Where $j = 1,2,\cdots,2t$.

Suppose $e(x)$ has *m* flips at locations $i_1, i_2, \cdots, i_m$. Then

$$e(x) = x^{i_1} + x^{i_2} + \cdots x^{i_m}$$

Where $0 \leq i_1 < i_2 < \cdots < i_m < n$

$S_1 = \eta^{i_1} + \eta^{i_2} + \cdots + \eta^{i_m}$

$S_2 = (\eta^{i_1})^2 + (\eta^{i_2})^2 + \cdots + (\eta^{i_m})^2$

$S_{2t} = (\eta^{i_1})^{2t} + (\eta^{i_2})^{2t} + \cdots + (\eta^{i_m})^{2t}$

For decoding purpose, we must find unknown values $\eta^{i_1}, \eta^{i_2}, \cdots, \eta^{i_m}$. Let us put $\eta^{i_l} = \rho_l$. Where $l = 1,2,3 \cdots, m$ then.

$$S_1 = (\rho_1) + (\rho_2) + \cdots + (\rho_m)$$

$$S_2 = (\rho_1)^2 + (\rho_2)^2 + \cdots + (\rho_m)^2$$

$$\vdots$$

$$S_{2t} = (\rho_1)^{2t} + (\rho_2)^{2t} + \cdots + (\rho_m)^{2t}.$$

Now error locator polynomial can be defined as

$$\begin{aligned}\sigma(x) &= (x + \rho_1)(x + \rho_2)(x + \rho_3) \cdots (x + \rho_m) \\ &= x^m + \alpha_1 x^{m-1} + \cdots + \alpha_m.\end{aligned}$$

Here $\rho = \{\rho_1, \rho_2, \cdots, \rho_m\}$ are the roots of polynomial $\sigma(x)$. To find these roots, we must calculate $\alpha_1, \alpha_2, \cdots, \alpha_m$. Where

$$\alpha_1 = \rho_1 + \rho_2 + \cdots + \rho_m$$

$$\alpha_2 = \rho_1\rho_2 + \rho_2\rho_3 + \cdots + \rho_{m-1}\rho_m$$

$$\vdots$$

$$\alpha_m = \rho_1\rho_2 \cdots \rho_m.$$

Now the relation between the syndrome and the coefficients of error locator polynomial is represented by Newtons Identities.

$$S_1 + \alpha_1 = 0$$

$$S_2 + \alpha_1 S_1 + 2\alpha_2 = 0$$

$$\vdots$$

For decoding of BCH-Code, the main task is to find error locator polynomial. For this purpose, we use a Berlekamp Massey algorithm. This is an iterative method and to start this method we need some initial conditions as follows.

π	$\sigma^{\pi}(x)$	d_{π}	l_{π}	$\pi - l_{\pi}$
-1	1	1	0	-1
0	1	S_1	0	0
1				
2				
.				
•				
2t				

Where d_{π} is known as an nth discrepancy,l_{π} is the degree of $\sigma^{\pi}(x)$,S_1 is the 1st non zero syndrome and t is the error correction capabilities.

To complete this table some steps are as under.

Step 1 If $d_{\pi} = 0$ then $\sigma^{\pi+1}(x)$=$\sigma^{\pi}(x)$ and $l_{\pi+1} = l_{\pi}$.

Step 2 If $d_{\pi} \neq 0$ then check $n < \pi$ such that $d_n \neq 0$ and $n - l_n$ has the greatest value in the last column then.

$\sigma^{\pi+1}(x) = \sigma^{\pi}(x) + d_{\pi} d_n^{-1} x^{(\pi-n)} \sigma^{n}(x)$ and $l_{\pi+1} = max(l_{\pi}, l_n + \pi - n)$.

Step 3 For discrepancy following formula is used

$$d_{\pi+1} = S_{\pi+2} + \sigma_1^{(\pi+1)} S_{\pi+1} + \cdots + \sigma_{l_{\pi+1}}^{(\pi+1)} S_{\pi+2-l_{\pi+1}}.$$

Thus in the last row, we get the required error locator polynomial $\sigma^{2t}(x)$.

Then we compute the roots of this polynomial and also find the inverses of these roots. These inverses of roots represent error location number and the exponent of error location number tells the error position in the received code. After subtracting the error vector from receiving vector we get correct code word.

2.3 Application of Coding Theory in Steganography

Coding theory is a branch of mathematics which deals with the problems of errors that happens to the message when it is transferred through some communication channels. The goal of coding theory is to offer secure communication of information, in the sense that errors that occurred during the transmission can be corrected. However, to enjoy the benefit of this property some value must be paid, this value is in the form of redundant bits which are added in the transmitting data. There are two purposes of utilizing error correcting codes in steganography. The first purpose is to secure the secret information which is loaded in some other medium from different attacks such as re-encoding and compression. The second purpose of utilizing error correcting code is to use syndrome coding for data embedding. The purpose of using syndrome coding for data embedding is to find those bits in the cover medium such that there is a minimum distortion in the cover medium. Here is some related work regarding steganography and coding theory.

2.3.1 DCT Based Video Steganography Using BCH-Code

In this work R J Mastafa utilize the concept of Discrete Cosine Transformation for video steganography along with BCH-Code. For the increase in security of secret information, he encrypted the message first by using a key and then encoded with BCH-Code. On the other hand, the video is taken as a cover medium. The video is first converted into frames and then each frame is separated into Y, U, V components. The purpose of converting frames into Y, U, V space is to remove the correlation between colors, then the encoded message is embedded in these spaces. He uses BCH-Code prior to the embedding due to the motivation from the work of Liu et al.

The proposed work of R J Mastafa consists of the following steps.

Input: video, a secret key for encryption, secret data.

Output: Stego video.

Step 1 Read video and separate into frames.

Step 2 Convert each frame into Y, U, V planes.

Step 3 Apply 2D-DCT on each plane.

Step 4 Read the message and encrypt using a key.

Step 5 Encode encrypted message using BCH $(7,4,1)$ Code.

Step 6 Convert encoded message into unit 8 and embedded in the frequency coefficients of each plain except DC coefficients.

The relation for embedding is shown below.

$$y_{ij}^{*} = \begin{cases} embed(y_{ij}, m_k) & ; y_{ij} \geq 0 \\ embed(|y_{ij}|, m_k) & ; y_{ij} < 0 \end{cases}$$

$$u_{ij}^{*} = \begin{cases} embed(u_{ij}, m_k) & ; u_{ij} \geq 0 \\ embed(|u_{ij}|, m_k) & ; u_{ij} < 0 \end{cases}$$

$$v_{ij}^{*} = \begin{cases} embed(v_{ij}, m_k) & ; v_{ij} \geq 0 \\ embed(|v_{ij}|, m_k) & ; v_{ij} < 0 \end{cases}$$

Where y_{ij}, u_{ij}, v_{ij} denotes the DCT coefficients of Y,U,V planes and m_k is the encoded message.

Block diagram of the embedding process is shown below.

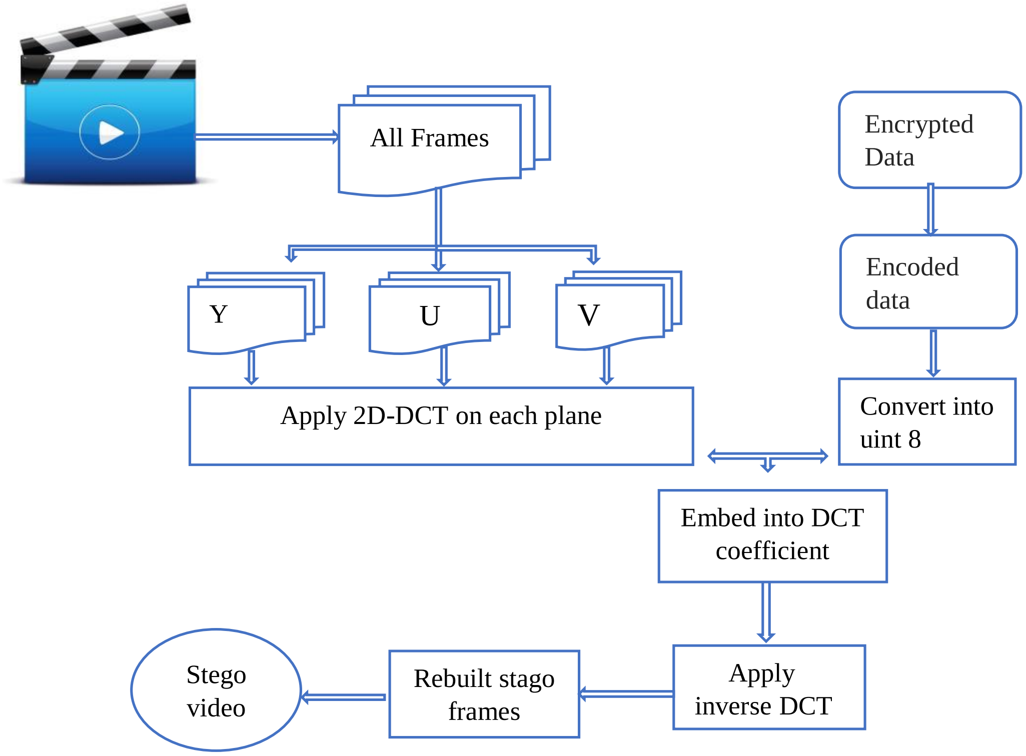


Figure 9 Block Diagram of Embedding Process

Extracting secret data from the stego video is done by using the reverse steps

Input: Stego video, a secret key for encryption.

Output: Secret data.

Step 1 Read stego video and separate into frames.

Step 2 Convert each frame into Y, U, V planes and apply 2D-DCT on each plane.

Step 3 Collect secret information by using a reverse relation as used for embedding purpose.

Step 4 Encrypt the extracted information by using a secret key.

For the experiment they utilized six uncompressed video format having a resolution (352×288). Some analysis is performed on these stego videos and concluded that the PSNR value of all the videos lies between 38 to 42dB.

Literature review

In 2014 Ramadhan J Mstafa and Khalad M Elleithy presented a video steganographic algorithm they used uncompressed video sequences as a cover media. Before embedding the secret message is encoded first using Hamming Code then this encoded message is embedded into the cover video frames [4]. In 2016 again, they presented a Discrete Cosine Transform (DCT) based video steganographic algorithm in this algorithm they used video as a cover medium. The secret message is first encoded using BCH Code then the encoded message is embedded into DCT coefficients of each Y, U, V planes except DC coefficients [2]. In 2012 Rengyue Zhang, Vasily Sachnev, Hyoung Jeong Kim and Jun Heo proposed a new data hiding scheme using BCH Code. Their method hides data inside a cover medium by modifying some coefficients of cover medium to null the syndrome. The proposed algorithm hides data with less computational time and less storage capacity [22]. In 2014 Fredrick R Lshengoma presented a new technique for steganography. He uses Reed Solomon Code for the steganographic purpose. According to his work, redundant bits of the R S Code can be replaced with the secret message bits [23].

Chapter 3

Video Steganography Using BCH-Code

There are two ways of securing information that is Cryptography and Steganography. The aim of Cryptography is to preserve the content of data secret, but using Steganography the existence of data can be made secret. Using Cryptography secret data can be encrypted such that an intruder can get approach to encrypted data. On the other hand, using a steganographic approach it may not easy for the intruder to find the existence of data. Therefore, the hybrid influence of both can make data much secure than the individual influence [10].

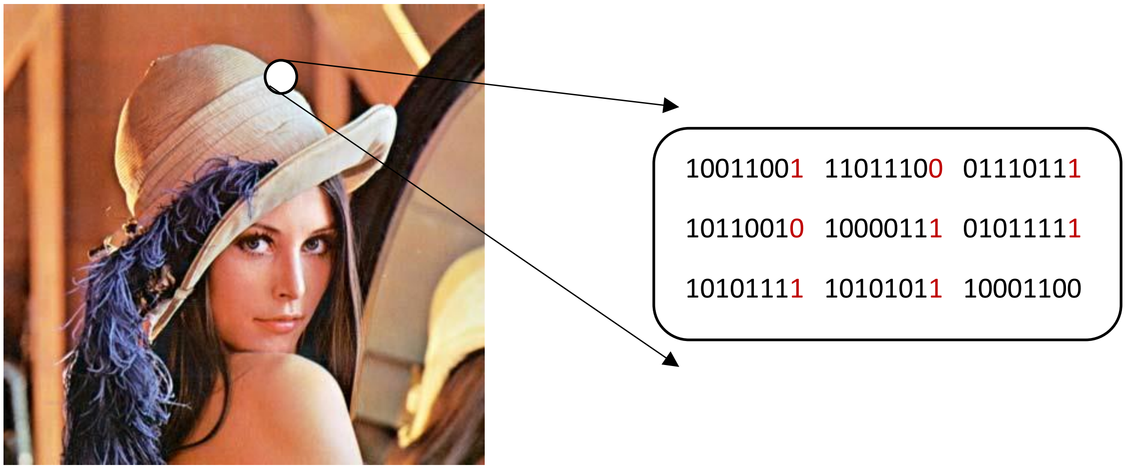
Steganography is the art and science of hiding information inside the cover medium. The cover medium may be Audio, Video, Text, and Images [5] Due to the extensive use of videos on the internet and having a large amount of redundancy it can be used as a cover medium [23]. Since the video is a collection of images so there are two main techniques used for image steganography (1) Spatial Domain technique (2) Transform Domain technique. In spatial domain technique data is embedded in the least Significant Bit and in Transform Domain technique data is embedded in the frequency coefficients of cover medium [3].

For the successful steganographic algorithm, two factors should be kept in mind. Which are embedding efficiency and embedding payload. The steganographic algorithm has a high embedding efficiency if the quality of stego data is good that is close to the original cover data so that the probability of data being hacked is very less. Embedding payload is the amount of data that is going to be hidden inside a cover medium. Both embedding payload and embedding efficiency are inverses of each other that means if we want high embedding efficiency we must decrease embedding payload. A lot of literature can be seen to balance embedding efficiency and embedding payload. According to the need for data security, embedding efficiency is more important. Therefore, to overcome the chance of data loss due to different attacks, we have increased embedding efficiency by using an LSB technique along with BCH-Code and pseudo-random number for frame selection further, we have used parallel embedding instead of sequentially to reduce the run time of the algorithm.

3.1 Least Significant Bit (LSB)

An image is made up of pixels and pixels are a combination of basic colors Red, Green, Blue. Each basic color is represented by 8 bits. There are 256 shades of each color so in a single pixel, total shades of colors are above 16 million.

The LSB is a common simple method for embedding information in a cover image. In this method least, significant bit or 8th bit of some or all bytes inside a pixel is changed with the bit of the secret message. A 24-bit pixel can store 3 bits of the secret message, thus 800 × 600 pixel image can store 180,000 bytes of the secret message. Consider 3 pixels of 24-bit Lena image and a character A whose ASCII value is (10101111) is embedded in these pixels as follows.



10011001	11011100	01110111
10110010	10000111	01011111
10101111	10101011	10001100

Then approximately half of the bits need to be modified so on average very small changes are required to embed secret messages these small changes cannot be detected by human eyes.

3.2 Proposed Method

In this section new algorithm for video steganography is proposed which uses the concept of pseudo-random number for frame selection symmetric cryptography for encryption BCH-Code for encoding and LSB for data embedding. The whole process of embedding consists of different phases.

3.2.1 Embedding Process

In the first phase, the video is converted into frames and some of these frames are selected using a pseudo-random number. These selected frames will be used as a cover medium for secret information. In the second phase whole, secret data are converted into equal blocks which is suitable for embedding. In the third phase encryption and the encoding process is completed.

Encryption of these data blocks is done with the help of a symmetric key algorithm. These encrypted data blocks are then encoded by using BCH-Code. Purpose of encoding data prior to embedding is to cure our secret information of different attacks such as re-encoding and compression. In the fourth phase, the encoded data are embedded in each selected frame using the LSB method. The architect of the proposed algorithm is shown in fig 8. The process of embedding is done using the following steps.

Input: Video, Seed, Secret information, the key for encryption.

Output: Stego-Video.

Step 1 Read video and select frames randomly.

Step 2 Read the secret information and divide into equal blocks.

Step 3 Secret information blocks are encrypted using symmetric key algorithms.

Step 4 Encrypted blocks of information are encoded using BCH-Code.

Step 5 Encoded blocks are embedded in each selected frame using LSB.

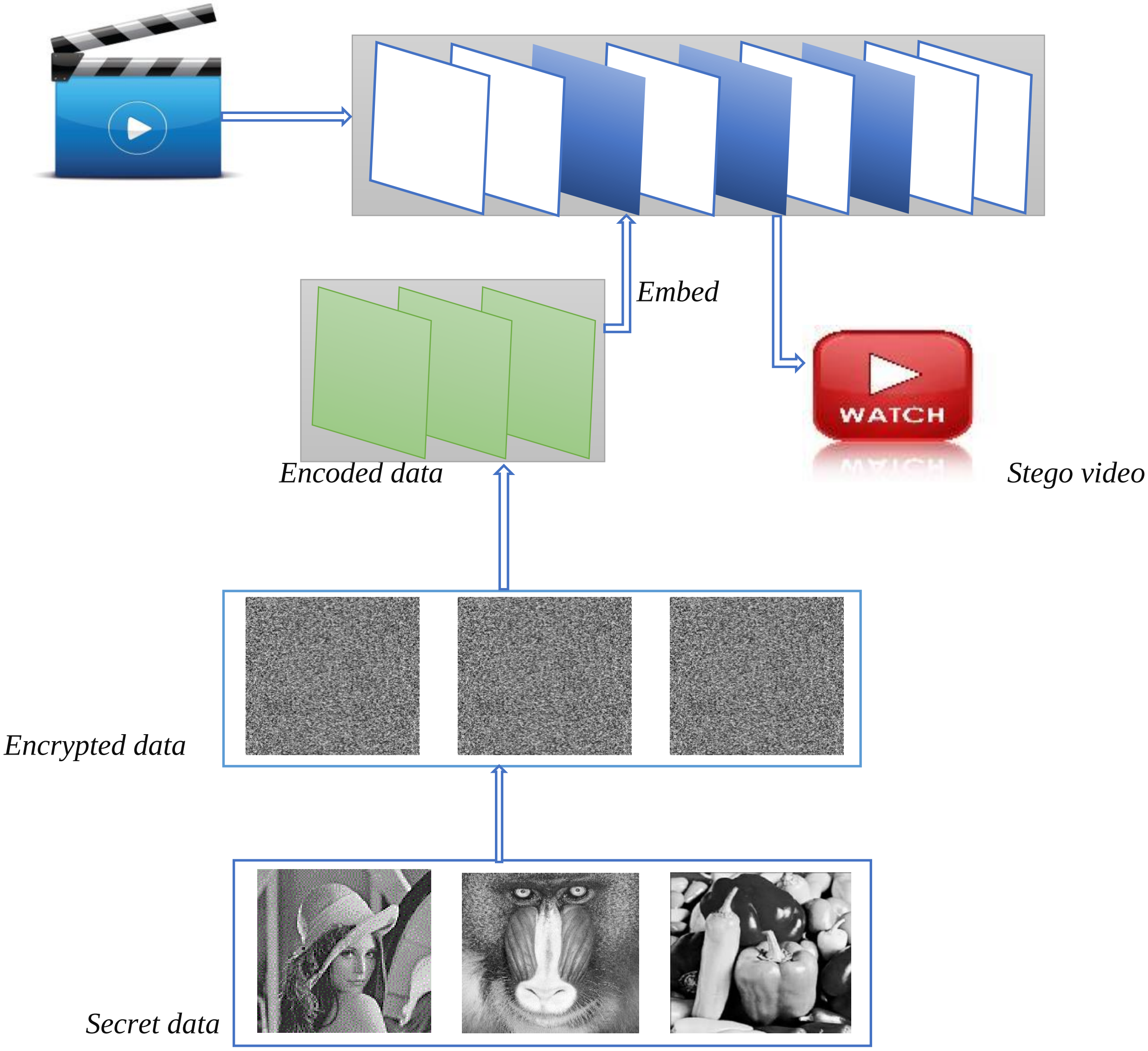


Figure 10 Embedding Process

3.2.2 Extracting Process

In data extracting process, the stego video is converted into frames and stego frames are selected using a pseudo-random number, then obtain hidden information from stego frames, after that this hidden information is decoded using BCH-Code thus we get encrypted data and this encrypted data is decrypted to get original data. Block diagram of the extracting process is shown in fig 11.

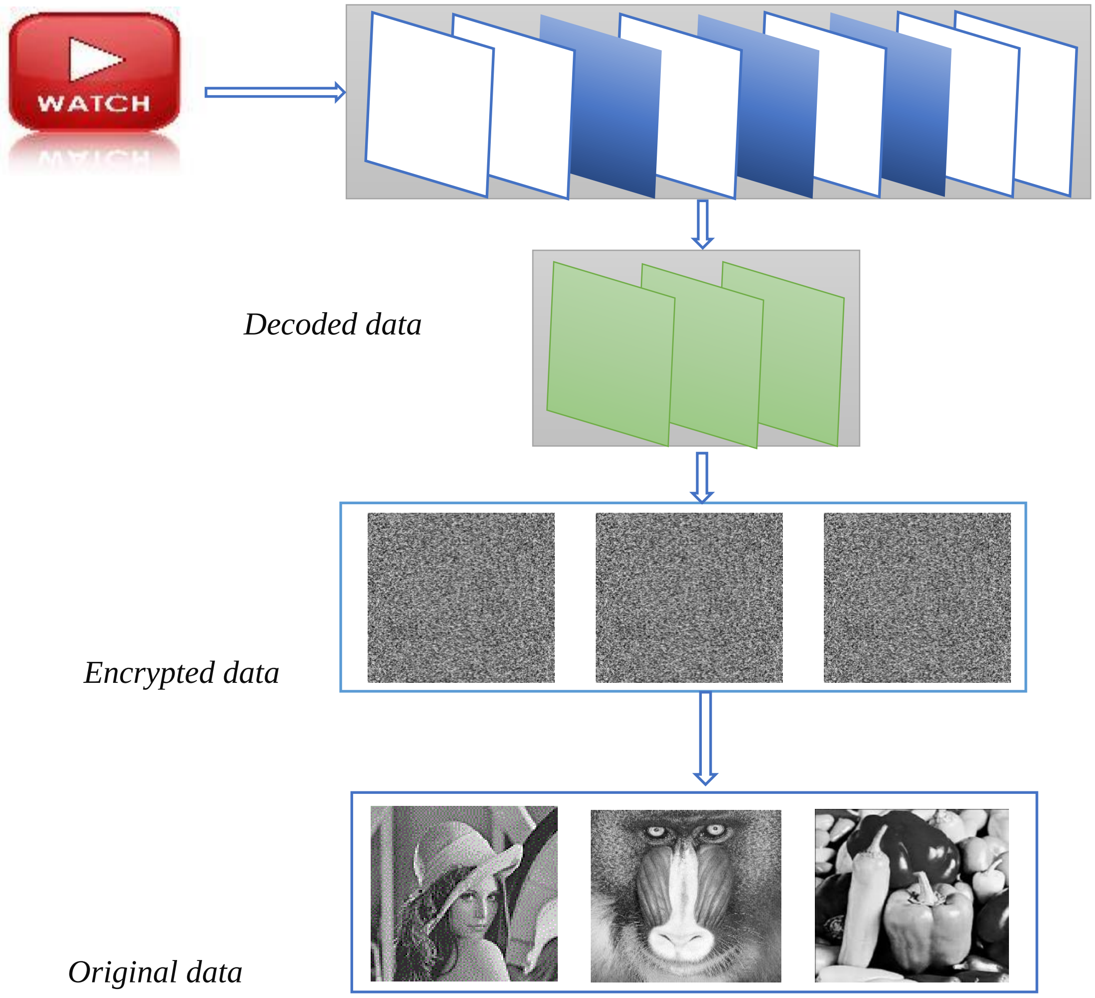


Figure 11 Extracting Process

Following steps are used to extract secret information.

Input: Stego-video, Secret key, Seed to generate PSRN.

Output: Secret information.

Step 1 Read stego video and collect stego frames.

Step 2 Extract secret information from these selected frames.

Step 3 Decode extracted data.

Step 4 Decrypt data using a secret key and get original information.

3.3 Experimental Results and Analysis

The proposed algorithm is tested on an uncompressed video stream with 300 frames of dimension 320×240. Five different secret images of equal size are selected as a secret information. Quality of stego frames is calculated from the PSNR value.

3.3.1 Mean Square Error (MSE)

The mean square error is used to check the consistency of the algorithm. It is calculated between stego and original image with the help of the following equation.

$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (O(i,j) - S(i,j))^2$$

Where $M \times N$ represents the size of image and $O(i,j), S(i,j)$ represents the pixel location at ith row and jth column of original and stego image respectively.

3.3.2 Peak Signal to Noise Ratio (PSNR)

To calculate the visual quality of the image PSNR is calculated. With the help of the following equation, we can calculate the PSNR of the image.

$$PSNR = 10log_2 \left(\frac{f_{max}^2}{MSE} \right)$$

Where f_{max} is the maximum pixel value of an image and MSE is the mean square error of the image. PSNR value is measured in dB if PSNR value is greater than 30dB then it will be difficult for the intruder to find the difference between original and stego image.

Remarks: For the experiment 5 cover frames from the video are selected which are 237, 76, 213, 186, 12, the Secret image of Lena is embedded in first selected frame Baboon is embedded in second selected frame Pepper is embedded in third selected frame Boat in fourth and Cameraman in fifth selected frames. The PSNR value of all the stego frames is listed in table 1. Further histogram of both cover and stego frames shows that there is no much difference between cover and stego frames as shown in figure:10 and figure 11 respectively.

Cover Frames	**Secret images**	**PSNR**	**R plain**	**G plain**	**B plain**
Frame237	Lena	47.66	50.63	52.00	44.35
Fraame76	Baboon	49.20	54.52	54.18	45.42
Frame213	Pepper	49.08	54.32	53.98	45.23
Frame168	Boat	46.67	49.46	49.34	42.25
Frame12	Cameraman	49.03	54.53	54.22	45.35

Table 1 PSNR and MSE value of Stego frames

Histogram of frame237 and stego frame 237 with the Lena image embedded in it are shown below

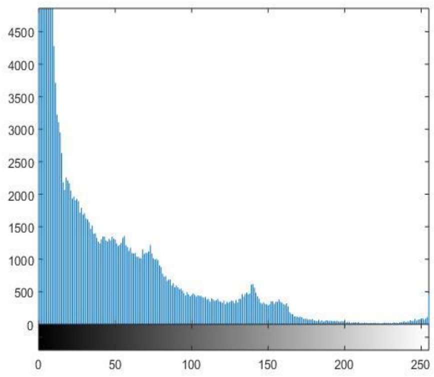


Figure 11 Histogram of Original Image

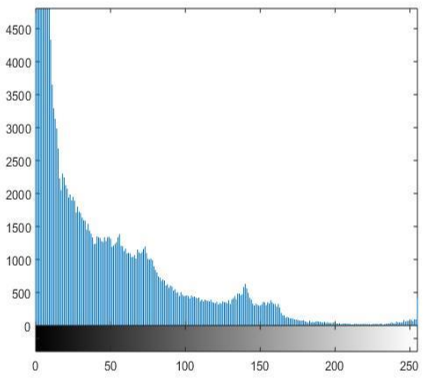


Figure 12 Histogram of Stego Image.

It is noted that the proposed algorithm is robust against different attacks. The first level of security is achieved by utilizing a symmetric cryptographic algorithm for encryption. The second level of security is attained by using a pseudo-random number for frame selection. The third level of security is reached by using BCH-(7, 4, 1) code. The fourth level of security is attained by embedding secret information in video frames. In the whole process, 2 keys are used first is the seed for frame selection and the second key is used for encryption. These two keys are sent to the receiver by using a secure channel.

Conclusion

In this work, we have presented an efficient technique for video steganography, which is based on BCH-Code and a pseudo-random number. In this algorithm, the video is divided into frames and selected frames are utilized for the hiding secret information. Many security levels are achieved in this algorithm. Because of the less modification in the cover frame, high embedding efficiency is achieved. Science high embedding efficiency claims the high visual quality of the loaded frame, so it is difficult for the hackers to suspect on the stego medium. Moreover, due to random frame selection, it is difficult to find the loaded frame with secret information. The security level of the proposed algorithm is also increased due to the influence of encryption and utilization of BCH-Code prior to the embedding of secret information. BCH-Code secure important information from reencoding and requantization. In the future work, we will use video as a secret information and embed in the video using the same process.

References

1. Al-Shatnawi, A. M. (2012). A new method in image steganography with improved image quality. Applied Mathematical Sciences, 6(79), 3907-3915.
2. Bansal, D., & Chhikara, R. (2014). An improved DCT based steganography technique. International Journal of Computer Applications, 102(14).
3. Bodhak, P. V., & Gunjal, B. L. (2012). Improved protection in video steganography using dct & lsb. International journal of engineering and innovative technology (IJEIT), 1(4), 31-37.
4. Brassil, J. T., Low, S., Maxemchuk, N. F., & O'Gorman, L. (1995). Electronic marking and identification techniques to discourage document copying. IEEE Journal on Selected Areas in Communications, 13(8), 1495-1504.
5. Channalli, S., & Jadhav, A. (2009). Steganography an art of hiding data. arXiv preprint arXiv:0912.2319.
6. Choudhury, M. R., & Bandyopadhyay, S. K. (2015). LSB Based Audio Steganography Using Pattern Matching. scanning, 2(11).
7. Cox, I., Miller, M., Bloom, J., Fridrich, J., & Kalker, T. (2007). Digital watermarking and steganography: Morgan kaufmann.
8. Baghel, R., & Sharma, P. (2012, August). Increasing robustness of LSB audio steganography using a novel embedding method. In National Conference on Security Issues in Network Technologies (NCSI-2012) (Vol. 5). Gwalior: IEEE Press.
9. Dhobale, D., & Ghorpade, V. R. (2013). An overview of advanced network protocol steganography. Int. J. Adv. Res. Comput. Communiaction Eng, 2(9), 2-5.
10. Ekatpure, P. R., & Benkar, R. N. (2013). A Comparative Study of Steganography & Cryptography. International Journal of Science and Research (IJSR), 4(7), 670-672.
11. Eltahir, M. E., Kiah, L. M., Zaidan, B. B., & Zaidan, A. A. (2009, April). High rate video streaming steganography. In Information Management and Engineering, 2009. ICIME'09. International Conference on (pp. 550-553). IEEE.
12. Gunjal, M., & Jha, J. (2014). Image Steganography Using Discrete Cosine Transform (DCT) and Blowfish Algorithm. International Journal of Computer Trends and Technology (IJCTT)–volume, 11, 144-150.

13. Ishengoma, F. R. (2014). The Art of Data Hiding with Reed-Solomon Error Correcting Codes. arXiv preprint arXiv:1411.4790.
14. KADHEM, S. M., ALI, M., & DHURGHAM, W. (2017). PROPOSED HYBRID METHOD TO HIDE INFORMATION IN ARABIC TEXT. Journal of Theoretical & Applied Information Technology, 95(7).
15. Karim, S. M., Rahman, M. S., & Hossain, M. I. (2011). A new approach for LSB based image steganography using secret key. Paper presented at the Computer and Information Technology (ICCIT), 2011 14th International Conference on.
16. Krishnan, R. B., Thandra, P. K., & Baba, M. S. (2017, March). An overview of text steganography. In Signal Processing, Communication and Networking (ICSCN), 2017 Fourth International Conference on (pp. 1-6). IEEE.
17. Liu, Y., Li, Z., Ma, X., & Liu, J. (2013). A robust data hiding algorithm for H. 264/AVC video streams. Journal of Systems and Software, 86(8), 2174-2183.
18. Mstafa, R. J., & Elleithy, K. M. (2014). A highly secure video steganography using Hamming code (7, 4). Paper presented at the Systems, Applications and Technology Conference (LISAT), 2014 IEEE Long Island.
19. Mstafa, R. J., & Elleithy, K. M. (2016). A DCT-based robust video steganographic method using BCH error correcting codes. Paper presented at the Systems, Applications and Technology Conference (LISAT), 2016 IEEE Long Island.
20. Paar, C., & Pelzl, J. (2009). Understanding cryptography: a textbook for students and practitioners: Springer Science & Business Media.
21. Sharp, T. (2001). An implementation of key-based digital signal steganography. Paper presented at the International Workshop on Information Hiding.
22. Simmons, G. J. (1984). The prisoners' problem and the subliminal channel. Paper presented at the Advances in Cryptology.
23. Singh, K. U. (2014). Video steganography: text hiding in video by LSB substitution. International Journal of Engineering Research and Applications, 4(5), 105-108.
24. Singh, S., Singh, E. S., & Sabo, T. (2015). IMAGE COMPRESSION USING DISCRETE COSINE TRANSFORM. IMAGE, 4(9).

25. Sudeepa, K., Raju, K., HS, R. K., & Aithal, G. (2016). A New Approach for Video Steganography Based on Randomization and Parallelization. Procedia Computer Science, 78, 483-490.
26. Tacticus, A. (1990). How to Survive Under Siege/Aineias the Tactician (Clarendon Ancient History Series): Oxford, UK: Clarendon.
27. Wilkins, E. H. (1954). A history of Italian literature: Harvard University Press.
28. Zhang, R., Sachnev, V., Botnan, M. B., Kim, H. J., & Heo, J. (2012). An efficient embedder for BCH coding for steganography. IEEE Transactions on information theory, 58(12), 7272-7279.
29. Fraleigh, J. B. (2003). A first course in abstract algebra. Pearson Education India.
30. De Andrade, A. A., & Palazzo Jr, R. (1999). Construction and decoding of BCH codes over finite commutative rings. Linear Algebra and Its Applications, *286*(1-3), 69-85.
31. Ferguson, N., & Schneier, B. (2003). Introduction to cryptography: principles and applications. NY: Wiley.