Hide a Secret File in Several BMP Images using the Circular Secret Key

Aamer Tahseen SUHAIL
{aamir@ntu.edu.iq, aamertahseen2017@gmail.com}

Ninaveh Technical Institute, Northern Technical University, Mosul, Iraq

Abstract: Digital data hiding technique is very important in achieving important and sensitive data security, especially those are transmitted through various digital communication channels, it differs from encryption technique, by unperceived or perceived by hackers, as it is based on the principle of hiding those Data within other digital media as their carrier covers. With the development of these hiding techniques, another technique developed to decipher the content of these data, to increase the efficiency of hiding methods of steganography and cryptography techniques should be merged. In this research the distribution of data to be hidden on a set of digital images of the type of BMP was adopted in a random way based on a circular secret key that created randomly and at an agreed length between the sender and the receiver, therefore this method eliminates the suspicion of any confidential content, as well as complicates the data retrieved process.

keywords: Steganography, encryption, circular secret key, cover, information hiding, least significant bit (LSB).

1 Introduction

The Internet and communication networks are one of the most important achievements in the field of information technology and data transmission, which require security. Encryption has been used as a technique and method to achieve this security of information, especially the important types that transmitted through various communication channels [1]. Various methods and algorithms have emerged to achieved encryption and decryption, the purpose of the encryption were to preserve the content of the messages confidential, and at the same time it was necessary to maintain the form of such messages to remove suspicion of piracy or hackers and this is not achieved by encryption techniques, it was necessary to use a new technique to achieve this, the technique of coverage (Steganography) was invented to achieve that [2][3].

Steganography is the art and technique of invisible and unperceived messages. This technique is achieved by hiding information in other information that serves as a cover for it. This concealment achieves the confidentiality of data transmission. The word (Steganography) derived from the Greek word "STEGO" which means "cover" and "GRAFIA" which means “writing”, and thus is known as “covered writing” [1][3].

IMDC-SDSP 2020, June 28-30, Cyberspace

DOI 10.4108/eai.28-6-2020.2298162

Steganography and encryption are both of them to protect information from unauthorized persons to view it, but there is a difference between them in that the encryption does not eliminate the principle of doubt in the data sent, while the steganography achieves the principle of eliminating this suspicion. Therefore, the technique of steganography has become of great importance in achieving the security of important and sensitive digital data, especially those transmitted between senders and receivers through various communication channels [1][4][5].

With the rapid development of information technique, the principle of confidentiality and reliability in the data is becoming a major challenge, so that the hiding technique was developed to achieve this principle because it is based on the mechanism of hiding secret information within multimedia such as images, audio or text. The methods of hiding generally rely on the mechanism of hiding in the least significant bits such as the first bit of the bytes of multimedia files that can be used as covers of secret such information without causing any defect or distortion in those media [5][6].

With the advancement of hiding techniques, and because it became common after achieving the efficiency and the desired goals, the suspicions come to the hackers and unauthorized persons slanderers in any received media exchanged between the peoples, especially those data issued by the important locations such as security or governmental and financial institutions, therefore, it is not enough to use the technique of hiding data in pure methods as simple sequential hiding of the data within these media carriers, but it was necessary to find common methods and hybrid algorithms that combine of cryptography and steganography techniques to increase hiding efficiency and achieve higher durability for the methods used [6][7].

The most important requirements for the efficient coverage model are [4][8].

a- Unexpected: means that no change or distortion has occurred in the cover after the hiding process.

b- Capacity: Amount of hiding or the capacity of the largest number of bits that can be hidden with no distortion in the results and features of the cover after the operation.

c- Robustness: The Strength and durability of hiding with withe inability to retrieve their contents by the attackers, especially in case of doubt or intrusion on those media.

2 Different Types of Coverage

In general, the majority of formats for digital files can be used as covers, but the most suitable files that achieve the highest abundance of the amount of redundancy, which can be defined as the number of bits in the digital entity that can be manipulated and replaced without causing any Change or distortion can be easily detected. Image files and sound files are among the best media to achieve these requirements [2][7].

2.1. Hiding in Images:

The images are represented in the computer as a Two-Dimensional Array of digital values, each one represents the intensity of a pixel unit. These values are used by the Raster to display on the screen at the location represented by each of these values [9].

This process depends on the number of binary cells used to represent each unit, the most important of which is the representation of (8 bit) or (24bit: 3Bytes) which is the so-called true color representation where each of the bytes represents a color of the three main colors (red, green, blue), and maybe represented by one cell only. There are several formats of digital images, and the most suitable for hiding, is the Image files with BMP extension, so this research will deal with eight bits and later on the representation of 24 bit will be discussed in the conclusion and future work [6][8][10].

2.2. Hide in the Least Significant Bit (LSB):

Because the most common method that is compatible with the mechanism of steganography is the hide in the least significant bit because the manipulation of the content of this field of those media does not affect their characteristics and does not occur any distortion in its features [6][11].

3. Proposed Method

We elected BMP image file to hide despite its large size compared to other types such as JPEG files, because of the capacity of the hiding area compared to other types, and to bypass this problem of large size, the pressing mechanism is used by using one of the lossless compression methods before sending these files through social media. The previous methods relied on using the secret key to achieve the randomness of hiding within the image file to increase the efficiency of hiding, but the weakness in those methods is in case of doubt of image hiding and theft the secret key, then it will be easy to access the secret message and extract its content [8][12][13].

To increase the efficiency of the previous methods, a mechanism was adopted to merge these methods with the addition of the random distribution feature in several images using the circular secret key. The method of distributing data or secret messages in several BMP images was proposed, and sent as separate files, depending on the method) algorithm) agreed between the sender and the receiver. The method is based on the mechanism of selecting several BMP image files as carrier covers for important secret files, provided that the images bearing sizes are larger or equal to the minimum size of eight times the size of the message that is hidden in the case of the LSB. After that, the circular secret key will be generated.

3.1. Circular Secret Key

The circular secret key contains the number of selected images as covers. For example, if there are four images, then the numbers of that key will be four, namely the numbers of those images. Even if the numbers of the images are (1, 2, 3, 4), the circular secret key will be as shown in Figure 1. below.

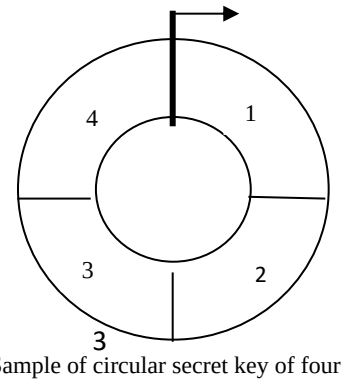


Fig. 1. Sample of circular secret key of four values

The four numbers are image numbers and bold font represents the starting point, and the direction of the arrow represents a trend of the application algorithm, thus the arrangement of the key numbers (1, 2, 3, 4), it will be on the process of hiding arrangement.

3.2. Characteristics of the circular secret key:

The efficiency of the proposed method depends on the efficiency of the circular secret key, and to increase the efficiency of this key must have the following characteristics:

a) The number of images in the key must be inserted randomly and not sequentially. For example, if the image numbers (1, 2, 3, 4), a random key arrangement is created as Figure 2.

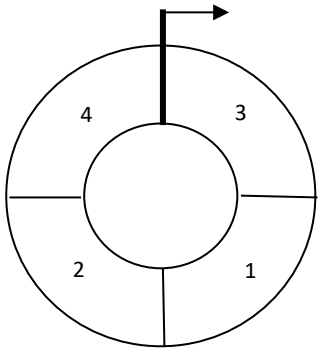


Fig. 2. Sample of circular secret key of four values in random distribution

The sequence of numbers is (3, 1, 2, 4).

b) To increase the efficiency of the key can be repeated images numbers on the key on one cycle in random sequence, this can increase the randomly of the message distribution, as well, as in Figure 3.

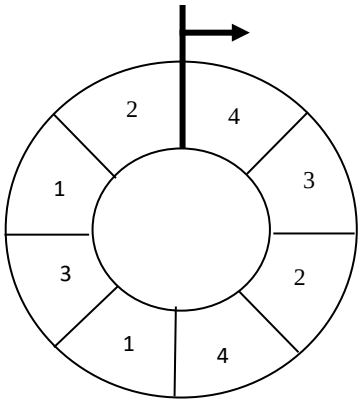


Fig. 3. Sample of the circular secret key of four values, the number of images repeated tow times randomly in one cycle

Thus, the images numbers are repeated twice in a single cycle and in a random way (4, 3, 2, 4, 1, 3, 1, 2), this increases the randomness of the hiding and distribution. The higher frequency of the numbers in a cycle, the greater the efficiency of the secret key

c) If the sizes of the selected images are different and not equal, the repetition percentage of the key numbers is not equal, and the percentage of repetition of each number depending on the size of its image to the sizes of the rest of the images, see Table. 1 below.

Table.1. Example of deferent images sizes

Image number	Image dimension in Pixel	Image size In Byte	Percentage of image size to the total size	Number of image iteration in one cycle of the circular secret key
1	100x200	20,000	30.77%	4
2	100x100	10,000	15.38%	2
3	100x50	5,000	7.70%	1
4	150x200	30,000	46.15%	6
total		65,000	100.00%	(key length)13

From the above, we find that the size of the secret key is 13 in length, and the numbers of the four images are repeated as the ratio of its size to the total size. The number of first images is repeated four times in one cycle of the secret key while the number of second images is repeated twice and the third one time, while the number of the fourth image is repeated six times during the cycle, and the distribution is random as previously indicated to increase efficiency see Figure 4. below, for example, the secret key can be as follows with the adoption of the random distribution of those numbers.

Image number	Number of image iterations
1	4
2	2
3	1
4	6

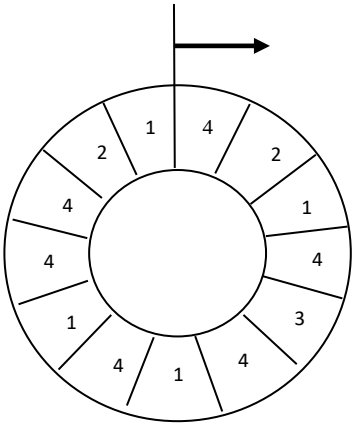


Fig. 4. Distribution of circular secret key of four different images sizes in one cycle

I.e., the circular secret key is (4, 2, 1, 4, 3, 4, 1, 4, 1, 4, 4, 2, 1).

d) The secret key can be sent in an independent form with a specific code to be agreed upon between the sender and the receiver. When the receiver receives that key, the number of images containing the hidden message will be known and then is distinguished from many of the images received based on a hidden code, determines being carrying a hiding message with its number.

3.3. Characteristics of Selected Covers

The selected images as cover caring hidden messages characterized by:

a) As the hiding takes place in the least significant bit of the image bytes, the percentage of hiding achieved will be 1/8 of the image size, then the size of the image (or the total size of the images in case of hiding in more than one image) should be equal to eight times the size of the data to be hidden at the least, see equations (1) and (2) in the following:

$$\sum_{i=1}^{i=n} (cov.\ [i]) >= (size\ of\ secret\ message*8) \quad (1)$$

or it can be checked as:

$$Size\ of\ secret\ message <= \sum_{i=1}^{i=n} (cov.[i])/8 \quad (2)$$

I.e.:
Cov. = cover image
i= image number
n= number of images

b) Not all of the selected images should be of the same size. They can be selected in different sizes provided that the above equation is achieved.
c) Hide a code that indicates that these images carry a hidden message, and hide the number of that image in certain locations of the image to be agreed between the sender and the receiver based on an agreed preset algorithm, to know which of the images carry a hidden message and the number of this image within the secret key, which will a guide to that key in hiding and extracting process. The images numbers not necessarily are sequential but it better to be random. The size of the space used to hide data must be subtracted from the total size of the image at calculating the hidden capacity of that image.
d) To increase the efficiency of the method, it is preferable to send images of hiding in different paths or at different times to prevent attempts in the case of discovery and to avoid the doubt.
e) The higher number of cover images lead to the higher efficiency of the method because the size of the circular secret key will be greater.
f) The number of bits allocated to represent the numbers of images does not have to be eight (one byte), but can be based on the number of images selected, for example, if the number of images is four, it is 2bits are enough to represent it as (00, 01, 10, 11), while if the number of 8 images we need 3bits to represent each one (000, 001, 010, 011, 100, 101, 110, 111). Therefore, the size of the secret key will be determined by the required bytes, so in the first example, it will be only one byte and in the second example it will be three bytes, and so on.

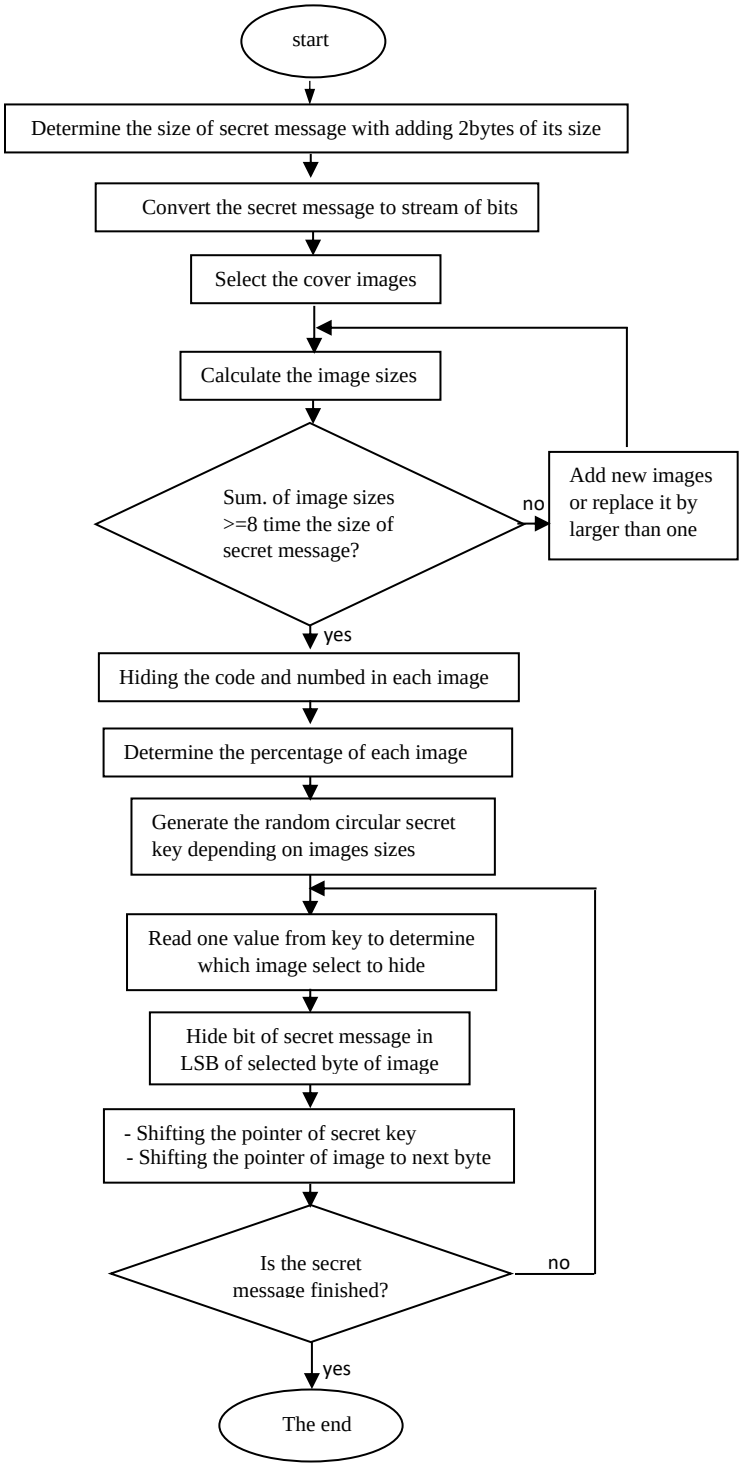
3.4. Secret message properties:

Because the proposed method depends on the mechanism of converting the confidential message file into a stream of bits before the hiding process, to be hidden in the form of sequential bits, it does not require a specific type of file types to hide, so it is valid for any type of digital file.

3.5. Algorithm of hiding:

After determining the images used as hiding covers and creating the circular secret key, the hiding process is performed by taking one number of circular secret key each time to indicate the image number in which a bit of the secret message to be hidden will be hide, after converting the secret message to a stream of bits and they are taken sequentially each time. When all circular key values are finished, the loop is repeated by rotating the process. Thus, all bits of confidential data are sequentially distributed in the selected LSB of selected image bytes based on circular secret key numbers. As detailed in the following algorithm:
Step 1: The Start.
Step 2: Determine the size of the secret message to hide, after adding two bytes in the beginning to represent the size of the hidden
message and secret code are agreed to indicate its end.

Step 3: Convert the secret message to a sequence of bits.
Step 4: Select the images to be used as covers.
Step 5: Calculate the total image sizes.
Step 6: Check whether images sizes are greater than or equal to eight times the size of a secret message if not, another image will be added or replaced with the larger ones.



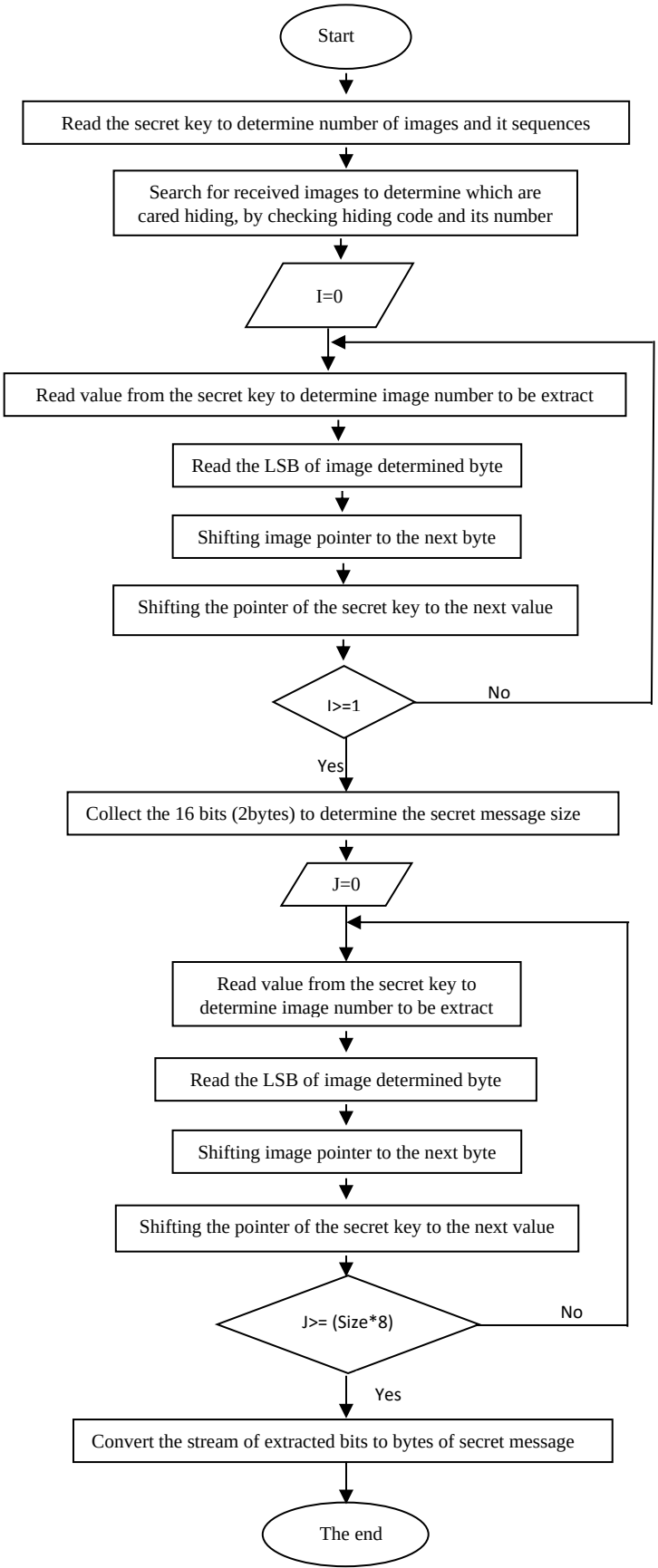
Step 7: Hide an agreed code between the sender and the receiver shows that the image is carrying a hidden message, as well as hide
the image number.
Step 8: Generate the circular secret key based on the numbers of images taking into account that each image should be repeated as
much as its size to the rest of the images.
Step 9: Read a value from the secret key and specify the number of the image to hide.
Step 10: Hide the secret message bit indicated in the least significant bit of the selected byte to hide from the selected image.
Step 11: Shifting the pointer of the image in which you hide is skipped.
Step 12: If all bits of the secret message has been hidden go to The end, otherwise: -
Step 13: Shifting the secret message pointer to the next bit.
Step 14: Shifting the secret key pointer to the next value.
Step 15: Go to step 9.
The end.

The detailed flowchart of the hiding algorithm is in Figure 5.

3.6. Algorithm of extract hiding:

The process of extracting the secret message from inside the images is an opposite process to what was done in the stage of hiding. After determining the dimensions of the images and numbers from reading the secret key, and identifying the images to hide from the collection of received images, which carries the symbol of hiding agreed between the sender and the receiver .The details of the hiding extracting process are shown in the following algorithm steps:

Step 1: The beginning.
Step 2: Reading the circular secret key to determining the numbers and sequence of images that are hidden.
Step 3: Identification of the images that contained the hiding from the collection of images received by checking the existence of the secret code, as well as determine the number of each image.
Step 4: Create a counter starting at zero to extract the 16 bits that contain the hidden secret message size.
Step 5: Reading a value of the circular secret key to determining the number of the image to be extracted from hiding.
Step 6: Read the least significant bit of the byte indicator from the image.
Step 7: Shift the current image pointer to the next byte.
Step 8: Increase the value of the counter by one.
Step 9: Did the value of the counter reaches 16, if so, the 16 bits extracted represent the size of the hidden file, otherwise refer to step 5.


Start
Read the secret key to determine number of images and it sequences
Search for received images to determine which are cared hiding, by checking hiding code and its number
I=0
Read value from the secret key to determine image number to be extract
Read the LSB of image determined byte
Shifting image pointer to the next byte
Shifting the pointer of the secret key to the next value
I>=1
No
Yes
Collect the 16 bits (2bytes) to determine the secret message size
J=0
Read value from the secret key to determine image number to be extract
Read the LSB of image determined byte
Shifting image pointer to the next byte
Shifting the pointer of the secret key to the next value
J>= (Size*8)
No
Yes
Convert the stream of extracted bits to bytes of secret message
The end

Step 8: Increase the value of the counter by one.
Step 9: Did the value of the counter reaches 16, if so, the 16 bits extracted represent the size of the hidden file, otherwise refer to step 5.
Step 10: Create a new counter starting at zero and ending with eight times the value extracted in step 9 which represents the size of the hidden file.
Step 11: Reading a value from the circular secret key to determining the number of the image to be extracted from hiding.
Step 12: Read the least significant bit of the byte pointer from the image.
Step 13: Shifting the current image pointer to the next byte.
Step 14: Increase the value of the new counter by one.
Step 15: If the value of the counter is eight times the size of the hidden file, then all the bits of the hidden file have been extracted, otherwise refer to step 11.
Step 16: Collect the bits of the extracted file by converting every eight bits to bytes to re-view the extracted file.
The end.
The detailed flowchart of the extract hiding algorithm is in Figure 6.

4. Conclusions and Future Work:

In the new method, the following conclusions and recommendations were reached.

4.1. Conclusions:

When applied this method to several cases that have shown their efficiency compared to the previous methods referred to, since it was difficult to reach the hidden secret message because of the following characteristics:

1- In the case of doubt about the existence of hiding, the process of extracting is very complex for the following reasons:

a) The extraction of the hidden message will produce random text, because what was hidden in it is part of several parts and in fragmented manner.
b) The numbers of images containing the hiding and their arrangement is unknown. so, no part of the secret message can be obtained in an of attempt to extract the hiding.
C) The theft of the secret key alone is not sufficient to reach the hidden text, it is impossible to know which of the images is hiding, as well as the random numbering and the secret location in which one was hidden.
2- The circular secret key achieves the randomness of the distribution, and the more random distribution of the numbers of images mean the more random there are, the same is true for increasing the repetition of those randomized in one cycle of that key. So that the efficiency of the method is greater.
3- The multiple images in which the hiding achieves a state of randomness as well as the difference in its size, which leads to inconsistency and regularity of the distribution of secret message bits, which increases the efficiency of the method by increasing random distribution.

4.2. Future Work:

1. In (24Bit) representation image, we can have more flexibility in dealing with this type of image. Each color can be considered as a separate image or more than one image, as well as the possibility of merging different types of images to increase the durability of the method
2. In the absence of a number of images as covers can be applied the method on one image by slicing it into several parts each part is treated as a separate image and is numbered randomly, then apply the same algorithm of hiding.
3. It is possible to combine several modes of hiding at the same time as the use of image files and audio files together, which increases the efficiency of the method and increase the complexity in extract hiding.
4. The percentage of hiding can be doubled by hiding in the 2 least significant bits, especially in those media where the change is not perceived as WAV sound files. Thus, the percentage of hiding is 1/4)25%).
5. To increase the efficiency of the method can encrypt the secret message with one of the encryption algorithms adopted before the process of hiding and this increases the random of hidden message in the case of trying to extracting it.

References

[1] N. Ch. Gowda, P. S. V. Srivastav, G. P. R.: Steg Cryp (Encryption using steganography). International Journal of Engineering and Advanced Technology (IJEAT), Vol. 8 (2019)
[2] Artz, D.: Digital Steganography: Hiding Data within Data", IEEE Internet Computing Journal (2001)
[3] Provos N. and Honeyman P.: Hide and seek: introduction to Steganography. IEEE Security and privacy Journal (2003)
[4] Zaidan, B. B., Zaidan, A. A. Al-Farajat A. K. and Jalab, H. A.: On the differences between hiding information and cryptography techniques: An overview. Journal of Applied Sciences (Faisalabad) 10 1650-1655 (2010)

[5] Kekra, H. B. Patankar A. B. and Koshti, D.: Performance Comparison of Simple Orthogonal Transforms and wavelet Transforms for Image Steganography. International Journal for Computer Applications, Vol. 44 (2012)
[6] Dewangga, I. G. A. P., Purboyo,T. W., Nugarachaeni R. A.: A New Approach of Data Hiding in BMP Image Using LSB Steganography and Caesar Vigenere Cipher Cryptography. International Journal of Applied Engineering Research, Vol.12 (2017)
[7] Douglas, M., Bailey, K., Leeney M., and Curran, K.: An overview of steganography techniques applied to the protection of biometric data. Multimedia tools and Applications, Vol. 77, 17333-17373 (2018)
[8] Karim, M., Rrahman, S, Hossain, I.: Anew approach for LSB based image steganography using secret key. in Proceedings of 14th International Conference on Computer and Information Technology. pp.-286-291 (2011)
[9] Bharti, P, and Soni, R.: A New Approach of Data Hiding in Images using Cryptography and Steganography. International Journal of Computer Applications, Vol. 58, (2012)
[10] Al- Husainy, M. A. F. and Uliyan, D. M.: Image Steganography Technique Based on Extracted Chains from the Secret Key. Journal of Engineering and Applied Sciences, Vol. 13, 4235-4244 (2018)
[11] Pillai, B., Mounika, M., Rao, P. J., and Sriram, P.: Image steganography method using K-means clustering and encryption techniques. In Advances in Computing, Communications and Informations (ICACCI), pp. 1206-1211 (2016)
[12] Karthikeyan, B., Kosaraju, A. C., and Gupta, S.: Enhanced security in steganography using encryption and quick response code. International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), pp. 2308-2312 (2016)
[13] Taha, M.S., Rahim, M. Sh. M. Iafta, S. A., Hashim M. M., and Alzuabidi, H. M.: Combination of Steganography and Cryptography: A short Survey. 2nd International Conference on Sustainable Engineering Techniques (ICSET) IOP Publishing Iop Conf. Series: Materials Science and Engineering Vol. 518 (2019)