Firefly Algorithm Based Data Hiding Technique

Vidya B¹, Bhagya Shree S R²

¹PG Student, ATME College of Engineering, Mysuru
²Associate professor, ATME College of Engineering, Mysuru

e-mail: vidyabpatel6646@gmail.com,srbhagyashree@yahoo.co.in

Abstract - In recent times a lot of work has been carried out in the field of reversible data hiding (RDH) to prevent the secret data from theft, illegal copying and unlawful reproduction. In RDH the cover image will be recovered after extracting the secret data which was embedded in that image. Finding the best location to hide the secret data is an important task so that it will conceal the existence of the message. This paper provides a reversible data hiding technique based on Firefly algorithm (FA). The optimal location to hide the secret data will be found by firefly algorithm. The histogram shifting technique is used to embed the secret data in the cover image. Histogram Techniques have attracted increasing interests due to their low computational complexity, high visual quality and can achieve good performance.

Key Words- Reversible Data Hiding, Firefly Algorithm, Histogram Shifting Technique.

1.INTRODUCTION

Multimedia content transmission over the internet creates plenty of security issues. These issues can be solved using the methods like, watermarking, cryptography and steganography. In watermarking process carrier signal is used to hide the digital information, the information that is hidden should, but it is not necessary to have a relation with carrier signal. At every distribution point, a digital signal is being embedded with a watermark. If a certain copy of the act is obtained later, then entire watermark can be retrieved and the distribution source is known easily. In steganography the complete secret message will be concealed into another different object which is known as cover object, it may be in image form, data format, audio / video. Cryptography is the technique, widely used for secured communication when there is an existence of third parties called adversaries. In general, cryptography is regarding establishing and examining the protocols that stop third parties or the community from reading or accessing private messages. It provides security but it is very complex method and occurrence of message can be obtained by the attackers using various decoding techniques.

Back Ground

A. Amsaveni and C. Arunkumar [1] proposed an efficient data hiding scheme using firefly algorithm in spatial domain, this paper provides a reversible data hiding technique based on Firefly algorithm (FA). The optimal location to hide the secret data will be found by firefly algorithm. The histogram shifting technique is used to embed the secret data in the cover image. Histogram Techniques have attracted increasing interests due to their low computational complexity, high visual quality and can achieve good performance. S. Picek and M. Golub [2] proposed a technique Evolutionary computation
algorithms represent a range of problem-solving techniques based on principles of biological evolution, like natural selection and genetic inheritance. Such algorithms can be used to solve a variety of difficult problems, among which are those from the area of cryptography. Md. Rashedul Islam, Ayasha Siddiqa, Md. Palash Uddin, Ashis Kumar Mandal and Md. Delowar Hossain [3] proposed a technique In Steganography, the total message will be invisible into a cover media such as text, audio, video, and image in which attackers don't have any idea about the original message that the media contain and which algorithm use to embed or extract it. Masoud Nosrati, Ronak Karimi and Mehdi Hariri [4] gave a detailed survey on one of data hiding techniques called "Reversible Data Hiding (RDH)". Hengfu YANG, Xingming SUN and Guang SUN [5] discussed Many existing steganographic methods hide more secret data into edged areas than smooth areas in the host image, which does not differentiate textures from edges and causes serious degradation in actual edge areas. Chi-Kwong Chan and L.M. Cheng [6] proposed a data hiding scheme by simple LSB substitution. Yongjian Hu, Heung-Kyu Lee, and Jianwei Li [7] narrates a difference-expansion (DE)-based reversible data hiding. Yun Q. Shi et al [8] discussed about Reversible Data Hiding in which the stego-media can be reversed to the original cover media exactly, has attracted increasing interests from the data hiding community. K. RamaKrishna and R. Sambasiva Rao [9] discussed about Swarm Intelligence (SI)-State-of-Art (SI-SA), Ziyad Tariq Mustafa, Bassim Abdul Baki Juma and Authman Waled Khalid [10] proposed Enhancing Protocol Steganography Using Firefly Algorithm.

2. OVERVIEW

Data hiding is a process which ensures exclusive data access to class members and protects object integrity by preventing unintended or intended changes. Data hiding also reduces system complexity for increased robustness. Data hiding is also known as data encapsulation or information hiding. Subsections describe about the three techniques in brief.

A. Watermarking

Watermarking or digital watermarking is the act of hiding a message or proprietary information related to a digital signal (i.e. an image, song, video) within the signal itself. Watermarking tries to hide a message related to the actual content of the digital signal, Digital watermarks may be used to verify the authenticity or integrity of the carrier signal or to show the identity of its owners. It is prominently used for tracing copyright infringements and for banknote authentication.

B. Steganography

Steganography is the art of hiding and transmitting data through apparently innocuous carriers in an effort to conceal the existence of the data, the word Steganography literally means covered or hiding writing as derived from Greek. Generally steganography is known as invisible communication. Steganography has its place in security. Hiding a message with Steganography methods reduces the chance of a message being detected. If the message is also encrypted then it provides another layer of protection.it do not alter the message structure but hides inside a cover object. Due to hidden or invisible factor it is difficult to recover hide information. Procedure to know the steganography technique is known as steganalysis.

C. Cryptography

The art of protecting information by transforming it (encrypting it) into an unreadable format, called cipher text. Only those who possess a secret key can decipher (or decrypt) the message into plain text. Encrypted messages can
sometimes be broken by cryptanalysis, also called code breaking, although modern cryptography techniques are virtually unbreakable. As the Internet and other forms of electronic communication become more prevalent, electronic security is becoming increasingly important. Cryptography is used to protect e-mail messages, credit card information, and corporate data. One of the most popular cryptography systems used on the Internet is Pretty Good Privacy because it's effective and free. Cryptography systems can be broadly classified into symmetric-key systems that use a single key that both the sender and recipient have, and public-key systems that use two keys, a public key known to everyone and a private key that only the recipient of messages uses.

The major factors for the information security is as follows:

Confidentiality: The secret information should not be seen by unauthorized person when the stego image is transferred.

Authentication: The received stego image should ensure that the image is received from authorized person.

Integrity: The received image should ensure that there is no alteration in the content when it is transferred between the source and destination.

Non-repudiation: It is a security service that ensures that an entity cannot refuse the ownership of a previous commitment or an action.

D. Reversible Data Hiding

Reversible Data Hiding is a type of data hiding technique that requires the cover object to be exactly recovered back after extraction of the secret data. It is widely used in the fields like, military, remote sensing, medical and judicial sectors.

Reversible data hiding (RDH) has been intensively studied in the community of signal processing. Also referred as invertible or lossless data hiding, RDH is to embed a piece of information into a host signal to generate the marked one, from which the original signal can be exactly recovered after extracting the embedded data. The technique of RDH is useful in some sensitive applications where no permanent change is allowed on the host signal. To evaluate the performance of a RDH algorithm, the hiding rate and the marked image quality are important metrics. There exists a trade-off between them because increasing the hiding rate often causes more distortion in image content. To measure the distortion, the peak signal-to-noise ratio (PSNR) value of the marked image is often calculated. Generally speaking, direct modification of image histogram provides less embedding capacity. In contrast, the more recent algorithms manipulate the more centrally distributed prediction errors by exploiting the correlations between neighboring pixels so that less distortion is caused by data hiding.

Least Significant Bit Method

Every pixel in an image indicates a color and the each image is made up of pixels. The lower values of the pixel in a gray scale image signifies dark areas and the higher values signify light areas. So in order to adjust the shades of the image its values can be adjusted and 8 bits are required to represent these values.

Difference Expansion

Difference expansion (DE) is a promising high capacity reversible data embedding algorithm. This technique received more attention over the years because of its high efficiency and simplicity. This is useful to achieve a very high embedding capacity, while keeping the distortion low.

3. Firefly Algorithm

Firefly Algorithm (FA) was first developed by Xin-She Yang in late 2007 and 2008 at Cambridge University
which was based on the flashing patterns and behaviour of fireflies. In essence, FA uses the following three idealized assumptions:

1. All fireflies are unisexual, so that one firefly will be attracted by all other fireflies;
2. Attractiveness is proportional to fireflies brightness;
3. For any two fireflies, the less brighter one will be attracted by the more brighter one;
4. The brightness can decrease when distance between fireflies increases;
5. If there are no fireflies brighter than the given firefly, then it will randomly move.

A. Behaviour of Fireflies

The flashing light of fireflies is an amazing sight in the summer sky in the tropical and temperate regions. There are about two thousand firefly species, and most fireflies produce host and rhythmic flashes. The pattern of flashes is often unique for a particular species. The flashing light is produced by a process of bioluminescence, and the true functions of such signaling systems are still debating. However, two fundamental functions of such flashes are to attract mating partners (communication), and to attract potential prey. In addition, flashing may also serve as a protective warning mechanism. The rhythmic flash, the rate of flashing and the amount of time form part of the signal system that brings both sexes together. Females respond to a male’s unique pattern of flashing in the same species, while in some species such as photuris, female fireflies can mimic the mating flashing pattern of other species so as to lure and eat the male fireflies who may mistake the flashes as a potential suitable mate. We know that the light intensity at a particular distance \( r \) from the light source obeys the inverse square law. That is to say, the light intensity \( I \) decreases as the distance \( r \) increases in terms of \( I \propto 1/r^2 \). Furthermore, the air absorbs light which becomes weaker and weaker as the distance increases. These two combined factors make most fireflies visible only to a limited distance, usually several hundred meters at night, which is usually good enough for fireflies to communicate.

B. Attractiveness of Fireflies

In the firefly algorithm, there are two important issues the variation of light intensity and formulation of the attractiveness. For simplicity, we can always assume that the attractiveness of a firefly is determined by its brightness which in turn is associated with the encode objective function. In the simplest case for maximum optimization problems, the brightness \( I \) of a firefly at a particular location \( x \) can be chosen as \( I(x) \propto f(x) \). However, the attractiveness \( \beta \) is relative, it should be seen in the eyes of the beholder or judged by the other fireflies. Thus, it will vary with the distance \( r_{ij} \) between firefly \( i \) and firefly \( j \). In addition, light intensity decreases with the distance from its source, and light is also absorbed in the media, so we should allow the attractiveness to vary with the degree of absorption. In the simplest form, the light intensity \( I(r) \) varies according to the inverse square law \( I(r) = I_s/r^2 \) where \( I_s \) is the intensity at the source. For a given medium with a fixed light absorption coefficient \( \gamma \), the light intensity \( I \) varies with the distance \( r \). That is \( I = 10e^{-\gamma r} \), where 10 is the original light intensity. In order to avoid the singularity at \( r = 0 \) in the expression \( I_s/r^2 \), the combined effect of both the inverse square law and absorption can be approximated using the following Gaussian form.
Sometimes, we may need a function which decreases monotonically at a slower rate. In this case, we can use the following approximation

\[ I(r) = \frac{I_0}{1 + \gamma r^2} \]  
\[ \text{.....(3.2)} \]

At a shorter distance, the above two forms are essentially the same. This is because the series expansions about \( r = 0 \) are equivalent to each other up to the order of \( O(r^3) \).

\[ e^{-\gamma^2r^2} \approx 1 - \gamma^2 r^2 + \frac{1}{2} \gamma^4 r^4 + \ldots \]  
\[ \frac{1}{1 + \gamma^2 r^2} \approx 1 - \gamma^2 r^2 + \gamma^4 r^4 + \ldots \]  
\[ \text{.....(3.3)} \]

As a firefly’s attractiveness is proportional to the light intensity seen by adjacent fireflies, we can now define the attractiveness \( \beta \) of a firefly by

\[ \beta(r) = \beta_0 e^{-\gamma r^2} \]  
\[ \text{.....(3.4)} \]

where \( \beta_0 \) is the attractiveness at \( r = 0 \). As it is often faster to calculate \( 1/(1+r^2) \) than an exponential function, the above function, if necessary, can conveniently be replaced by

\[ \beta = \frac{\beta_0}{1 + \gamma r^2} \]

Equation 3.4 defines a characteristic distance \( \Gamma = 1/\sqrt{\gamma} \) over which the attractiveness changes significantly from \( \beta_0 \) to \( \beta_0 e^{-1} \).

In the implementation, the actual form of attractiveness function \( \beta(r) \) can be any monotonically decreasing functions such as the following generalized form

\[ \text{.....(3.5)} \]

For a fixed \( \gamma \), the characteristic length becomes

\[ \beta(r) = \beta_0 e^{-\gamma r^m}, \quad (m \geq 1) \]
\[ \Gamma = r^{-1/m} \rightarrow 1 \text{ as } m \rightarrow \infty. \]

Conversely, for a given length scale \( \Gamma \) in an optimization problem, the parameter \( \gamma \) can be used as a typical initial value. That is

\[ \gamma = \frac{1}{\Gamma^m} \]  
\[ \text{.....(3.6)} \]

C. Distance and Movement of Fireflies

The distance between any two fireflies \( i \) and \( j \) at \( x_i \) and \( x_j \), respectively, is the Cartesian distance

\[ r_{ij} = ||x_i - x_j|| = \sqrt{\sum_{k=1}^{d} (x_{i,k} - x_{j,k})^2} \]  
\[ \text{.....(3.7)} \]

where \( x_{i,k} \) is the \( k \)th component of the spatial coordinate \( x_i \) of \( i \)th firefly. In 2-D case, we have

\[ r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \]  
\[ \text{.....(3.8)} \]

The movement of a firefly \( i \) is attracted to another more attractive (brighter) firefly \( j \) is determined by

\[ x_i = x_i + \beta_0 e^{-\gamma r_{ij}^2} (x_j - x_i) + \alpha (\text{rand} - \frac{1}{2}) \]  
\[ \text{.....(3.9)} \]

where the second term is due to the attraction while the third term is randomization with \( \alpha \) being the randomization parameter. \( \text{rand} \) is a random number generator uniformly distributed in \([0, 1]\). For most cases in our implementation, we can take \( \beta_0 = 1 \) and \( \alpha \in [0, 1] \). Furthermore, the randomization term can easily be extended to a normal distribution \( N(0, 1) \) or other distributions. In addition, if the scales vary significantly in different dimensions such as \(-10^5\) to \(10^5\) in one dimension while, say, \(-0.001\) to \(0.01\) along the other, it is a good idea to replace \( \alpha \) by \( \alpha S_k \) where the scaling parameters \( S_k (k = 1, \ldots, d) \) in the \( d \) dimensions should be
determined by the actual scales of the problem of interest. The parameter $\gamma$ now characterizes the variation of the attractiveness, and its value is crucially important in determining the speed of the convergence and how the FA algorithm behaves. In theory, $\gamma \in [0, \infty)$, but in practice, $\gamma = 0(1)$ is determined by the characteristic length $\Gamma$ of the system to be optimized. Thus, in most applications, it typically varies from 0.01 to 100.

D. Importance of Firefly Algorithm

Firefly algorithm has attracted much attention and has been applied to many applications. It requires least computation time for digital image compression. It can be efficiently used for feature selection and showed it produces consistent and better performance in terms of time and optimality than other algorithms.

In the engineering design problems it can efficiently solve highly nonlinear, multimodal design problems. In antenna design optimization it can outperform artificial bee colony (ABC) algorithm. It can also outperform PSO and obtained global best results.

For discrete problems and combinatorial optimization, discrete versions of firefly algorithm have been developed with superior performance which can be used for travelling-salesman problems, graph coloring and other applications. In addition, extension of firefly algorithm to multi objective optimization has also been investigated A few studies show that chaos can enhance the performance of firefly algorithm while other studies have attempted to hybridize FA with other algorithms to enhance their performance.

FA has two major advantages over other algorithms: automatically subdivision and the ability of dealing with multimodality. First, FA is based on attraction and attractiveness decreases with distance. This leads to the fact that the whole population can automatically subdivide into subgroups, and each group can swarm around each mode or local optimum. Among all these modes, the best global solution can be found. Second, this subdivision allows the fireflies to be able to find all optima simultaneously if the population size is sufficiently higher than the number of modes. Mathematically, $1/\sqrt{\gamma}$ controls the average distance of a group of fireflies that can be seen by adjacent groups. Therefore, a whole population can subdivide into subgroups with a given, average distance. In the extreme case when $\gamma = 0$, the whole population will not subdivide. This automatic subdivision ability makes it particularly suitable for highly nonlinear, multimodal optimization problems.

In addition, the parameters in FA can be tuned to control the randomness as iterations proceed, so that convergence can also be speed up by tuning these parameters. These above advantages makes it flexible to deal with continuous problems, clustering and classifications, and combinatorial optimization as well.

E. Histogram Shifting Technique

In Histogram Shifting Technique the input image is partitioned into pieces or blocks after that histogram shifting is performed on every block which upgrades the data concealing capacity and visual quality. Quantity of data that could be embedded inside the blocks of image is greater on comparison with embedding inside a particular single image.

This RHD technique primarily comprises of four important phases:

1) Dividing an image into blocks
2) Processing Phase and
3) Embedding Phase
4) Extraction Phase.

First stage comprises of isolating the given image into blocks. Processing stage involves producing the histogram of each and every block and then taking the distinction of histogram once the histogram modification is done.

Ni et al uses the zero or least histogram point. If in case, the peak is lesser than the zero or lower point / value in the
histogram, it increments the value of the pixel by 1 from greater than the peak to lesser than the zero or minimal point in the histogram. During embedding process, the entire image will be searched. When a peak-pixel value is encountered, if the bit to be embedded is '1' the pixel is added by 1, else it is kept intact. Alternatively, if the peak is higher than the zero or minimum point in the histogram, the algorithm decreases pixel values by one from lower than the peak to higher than the zero or minimal factor within the histogram, and to embed bit '1' the encountered peak-pixel value is subtracted by 1. The interpreting procedure is quiet easy and reverse of the embedding method.

The advantages of this method are as follows:

• It is easy,
• Distortions are quite undetectable
• Capacity is high.

The disadvantage are as follows:

• Capacity is restricted by using the frequency of peak-pixel price within the histogram
• It seeks the picture a couple of instances, so the algorithm is period imperceptible.

F. Block Diagram of Proposed Method

Block diagrams shown in Figure 3.1 represents the proposed method for embedding secret data into the cover image. Each step is explained in the next sections.

Embedding the Secret Data

The steps to embed the key data into quilt image are presented below

Step 1: Get the cover image C of size MxN and the secret data S.

Step 2: Each individual character of the secret data is obtained.

Step 3: The cover image is resized into 256X256 non-overlapping blocks. It is denoted as B, data to be hidden in a block is denoted as r = L/n, where l = 1, 2,…., n.

Step 4: To find the best location in each block using firefly algorithm, we need to perform the following:

• Set the initial parameters: nf (no. of fireflies), its maximum count of iterations, randomization parameter (between 0|& 1), C1 and C2 (constants) rand (random number).

• Each firefly has the size of number bits to be hidden in each block r. If r is 5 then each firefly represents the group of 5 bits which is chosen from 64 pixel block.

• The firefly algorithm (FA) is an iterative algorithm. In each iteration, the stego image block for each firefly is obtained by embedding r bits of secret data in that firefly position. The best location is found when the following conditions occur:

  ▪ Number of iteration exceeds maximum number of iterations
  ▪ No development is obtained inside the successive iterations
  ▪ An acceptable result has been found

Step 5: Embedding Process: The r bits from secret data is embedded in the best pixel of each block. For this purpose the histogram shifting method is used.
For a gray scale image the pixel intensity value $x \in [0= 255]$. The histogram $H(x)$ is generated based on the equation 3.11.

$$H(x) = p_x, \ 1 \leq x \leq 255 \ \ \ \ \ \ \ \ \ \ (3.10)$$

- The maximum intensity pixel or peak point and minimum intensity pixel or zero point are found in the histogram $H(x)$.
  - If the peak point is greater than the 0 point, then all the pixel values in the histogram are decremented by 1. It means that the entire histogram is shifted towards the left by 1 unit.
  - The image is checked in the consecutive order. Once the peak point is met, then the secret bit is checked. If secret bit is one then the peak point value is incremented by 1. If secret bit is 0 then there is no change in the peak point value.

**Extracting the Secret Data**

There are three main steps to extract secret data, each one explained as follows:

Step 1: The stego image is scanned in the similar order which is second-hand in the embed process. If the maximum intensity pixel-value is greater than peak point then bit 1 is extracted and if the maximum intensity pixel value is equal to peak point then bit 0 is extracted.

Step 2: The value of maximum intensity pixel is decremented by 1 whenever the bit 1 is extracted.

Step 3: After extracting all bits the entire value of the histogram is incremented by 1. So that the empty place which is created during embedding process is removed.

**4. RESULTS**

This section describes about the implementation results of data hiding process into cover images and cover video through pictorial representation.

**A. Snapshots of Complete Integrated Module for Image**

The Graphical User Interface shown in the Figure 4.1 represents the steps implicated in the entire data hiding process into a cover image. Each step is explained with an example below,

**Figure 4.1: Graphical User Interface for Image**

**Step 1:** Initially the cover images has to be selected in which the data is to be hidden. And the image is resized into 256X256 order. The cover image must be a color image of size MXN. Figure 4.2 shows the two cover images considered.
Step 2: Enter the data which is to be hidden into the cover images. Here there is a provision for entering two data’s. Input data can be an alphabet, numerical or special characters. Each individual character of the input data obtains an optimal location in the considered cover images.

The best locations for embedding are found using firefly algorithm. Since number of best locations is equal to the numbers of characters in the secret message, N (i.e. Number of fireflies for the firefly algorithm) is taken as number of characters in the secret message. While taking the best locations there is chance of repeated locations considered as the best thereby making the embedding lossy. So in order to have lossless embedding the best location should be non-repetitive. So a care is taken in this regard to achieve the non-repeated best location in the design.

Step 3: Corresponding locations of each individual character is represented in terms of rows and columns i.e., (r,c). These locations are selected randomly and it is non repetitive also. Figure 3.3 depicts the data embedded locations.

Step 4: As we can enter two data’s, we can obtain the embedded locations of each data individually. Figure 4.4 and Figure 4.5 represents the optimal locations in the (r,c) format for Data-1 and Data-2 respectively.

Figure 4.2: Cover Images

Figure 4.3: Data Embedding Locations

Figure 4.4: Data Embedding Locations of Text-1
Step 5: Data Extraction Details is represented in the Figure 4.6. It indicates the extraction of original data without any modifications.

Step 6: Time Analysis is depicted in Figure 4.7. It provides the Embedding time and Extraction time required for Data-1 and Data-2 respectively.

Step 7: Stego images are those which contain the hidden data in the respective locations obtained from the Firefly Algorithm. Stego images are as shown in the Figure 4.8.

Step 8: This step closes all the matlab windows related to the above steps.

5. CONCLUSION AND FUTURE WORK

CONCLUSION

An efficient reversible data hiding scheme using Firefly Algorithm for hiding secret data into cover image has been proposed. The proposed method finds the optimal locations to hide the secret data using FA. The secret data is embedded in those locations via histogram shifting technique. It provides a high security as each character of the hidden data is located at different
positions. Two images are considered as cover images more different data's can be hidden into them separately. It allows the user to hide more amount of data. It requires less computation time and provides a consistent and better performance. It can be used in the fields of medical, military and judicial sector as it provides a better confidentiality.

FUTURE WORK

In the proposed method only two images are considered as cover images to hide the secret data. In future work it can be extended to n number of cover images such that a huge quantity of the data can be hidden.

6. REFERENCES


