Acta Scientiarum

http://periodicos.uem.br/ojs
ISSN on-line: 1807-8664
Doi: 10.4025/actascitechnol.v44i1.57581

SFO for enhancing steganogrphy by exploiting exact resemblance between cover and secure text

Omar Younis Abdulhammed

Department of Computer, College of Science, Garmian University, Kalar, Bardasure, Iraq. *Author for correspondence. E-mail: Omar.y@garmian.edu.krd, omaralaa78@yahoo.com

ABSTRACT. Steganography has become an important science in security in recent years, it hides the existence of the secret information in such a way that no one suspects the information exists. In this paper a new steganography method for concealing message in message by using sailfish optimizer (SFO) and transposition method is proposed. In this work, the (SFO) is utilized to enhance the security and reduce the disfigurement to obtain best message quality, where the SFO has advantages in the term of security and text quality. The fitness values of the cover and secret message are calculated by using a new way. The SFO works to determining the matching fitness value between the cover and secret message and generates a key to recapture secret message. Lastly, relying on the matching fitness value, the steganography process is done by replacing (transposition) the secret's letters with cover's letters. The results proved that proposed method is effective, efficient and provided high security, high capacity, not malformed text and resistance against several steganalytic attacks.

Keywords: Steganography; cover message; secret message; SFO; steganalytic.

Received on January 27, 2021.
Accepted on March 5, 2021.

Introduction

As long as message transmission on the internet still face some problems such as data security and copyright control, it's time to find out secret communication schemes for transmitting message on the internet safely (Maleki, Jalali, & Jahan, 2014). One of the most important techniques used to preserve information during transmission and receipt via the internet is data hiding. This technique which attempts to hide secret "messages" is considered very difficult to decode (Alwan, Farhan, & Mahdi, 2020). Information hiding generally involves several sub-disciplines in the information security field (Baawi, Mokhtar, & Sulaiman, 2017) such as watermarking, digital signatures, steganography and cryptography. Each of these methods has its own advantages and disadvantages. Watermarking is widely used for copyright protection; digital signatures are widely used for the protection of digital signatures [Setyono & Setiadi, 2019), while steganography and cryptography are widely used to secure digital messages (Setyono, Setiadi, & Muljono, 2017). The Cryptography scrambles the messages so it is not easily understood. Steganography diverges from cryptography, it conceal the existence of the embed information (Kaur & Rani, 2016).

There are different text steganography methods used by numerous investigators, but a little number of them concentrate on improve security, enhance visual quality and increasing payload capacity together. The main objective of this paper is to increase the security, payload capacity and maintain stego text quality of the text steganography by using new idea based on SFO and transposition method, where the some letters of the cover text are replaced (transposition) by the letters of the secret text by relying on the similarity between fitness values of the cover and secret text, whereas the SFO is doing the search for finding similarities between them, also a novel method was used to obtain the fitness values.

Steganography

One of the important principles in any communication process between the sender and the recipient via the internet is security. Although the management of threats and risks is possible by using modern security techniques but these technologies are not enough for maintain information security. Therefore, additional mechanisms of security are needed to secure information (Shukur & Jabbar, 2018). Steganography is an art of science for discrete conveyance; it is a peculiar method of shielding messages. Its objective is concealing secret information within a cover image to avoid visibility to an attacker with visibility to only the transmitter and beneficiary of the information who have message reality realization. The central principles of steganography entail a text, a carrier object, hiding technique and a stego key for improved security. The carrier object entails embedding information which ought to be text, audio or video. Employing steganography can happen in a wide span of approaches inclusive of reliable conveyance of secret military information and other intelligence organizations, strengthening mobile banking reliability, reliability of online voting, and conceal disclosure between two relaying parties (Abdulwahed, 2020). Steganography can be effectively used in different applications, but using it can be quite risky because attackers can use it to send Trojans and viruses with the aim of compromising sensitive systems. Moreover, with the use of this information hiding technology, criminals or terrorists may be enabled to exchange secrete information (Taha, Rahim, Lafta, Hashim, & Alzuabidi, 2019). However, there are three important parameters (criteria) in designing steganography systems, that is, security, robustness and capacity as shown in Figure (1) (Mohamed, 2014).

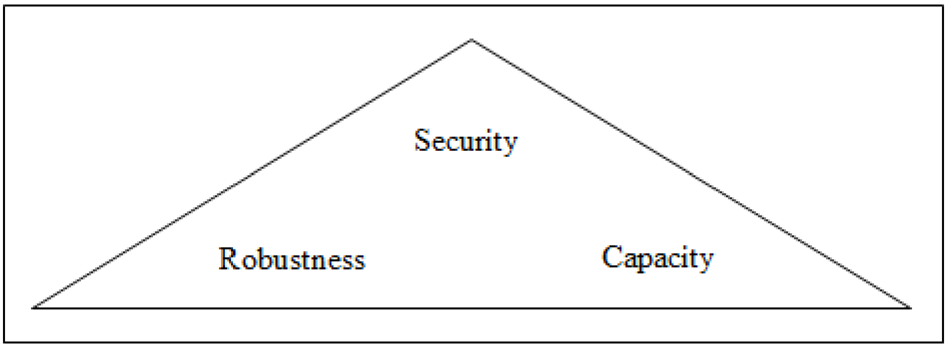


Figure 1. Steganography's Parameters.

Text steganography is the art or process of hiding a text into another to secure communication. It ensures that unauthorized users cannot obtain the secret message for their private use (Singh, Chaudhary, & Agarwal, 2012). As previously mentioned, text steganography is regarded as the most difficult technique compared with audio and image steganography because of insufficient redundant data, which is common in other carriers, rendering most of its techniques with insufficient capacity and security even though these techniques are generally transparent (secret) (Roy & Venkateswaran, 2013; Kingslin & Kavitha, 2015).

However, the ability of cover files (such as text files) to embed secret data depends on the availability of redundant or insignificant within them. The characteristics of a cover file that is changed, manipulated, or modified during the embedding process should remain invisible to unauthorized users. There are three main categories used to hide text-in-text messages, that is, format based, random and statistical generations, and linguistic method as shown in Figure 2.

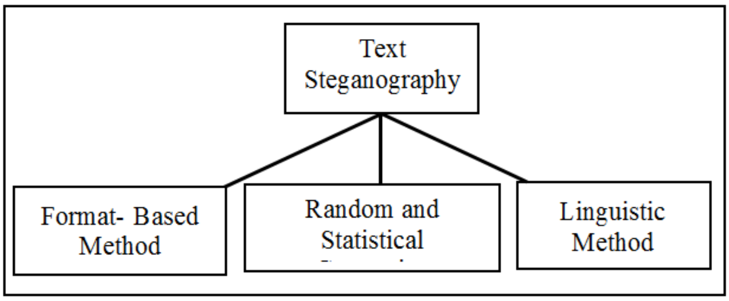


Figure 2. Categories of Text Steganography.

Format based

In this method, the covering message will not be changed in case of its words and sentences; modifications will just be made on the spaces between the words, lines, or/and paragraphs using special characters, that is, whitespace steganography. This method has many limitations. For example, the length of binary string must be less than or equal to the number of word spaces.

Random and statistical generation methods

These methods automatically generate the cover text message; it does not need an existing cover message. The generated cover message uses the secret message in generation process. This algorithm uses this language structure and properties—i.e., how to create the sentences, what is the past format of a verb … etc. Also, these methods use grammar to produce suitable cover message. In this type of text steganography, an extra complexity is added (time and space) to generate a full paragraph; this consumes long time to embed and extract the secret message in/from the cover message.

Linguistic methods

This method is used to hide a message in another message depending on the linguistic structure of the cover message (the punctuation marks) or the words semantic as a place to hide the message, message, which has two main types that are Linguistic structure and semantic method, This technique has drawbacks, such as the data size depends on the number of the punctuation marks of the cover message, and failed in protecting the sent message when the outsider tries to find the original message by swapping each word to the original one using the semantic algorithms (Hamdan & Hamarsheh, 2016). This paper introduces a new approach for text steganography by using transposition process that achieved by using sailfish algorithm, this approach has overcome all the disadvantages that have been shown in the previous methods.

The sailfish optimizer

Recently, Shadravan, Naji, and Bardsiri (2019) developed a new meta-heuristic called Sailfish Optimizer which combines the behavior of both a group of sailfish as a predator and a group of sardine as the prey. The sailfish is classified as a social predator because it is characterized by working in group to catch and hunts its prey. In a cooperative hunting, predators use different strategies to kill. For example the group of sailfish is characterized by the alternation of attacks strategies. It consists in that, at a given moment, each member of group attack alone the school of prey (sardine) and injure or hunt some of them, while the other members of group save their power. When a sailfish attacks the prey school he can update his position around them. In addition, the sailfish can also update his position to occupy empty space around the prey school and simulate encircling the prey. On the other hand, in order to escape from the following attacks of the sailfishes, the prey group (sardine) changes the position when a member of their group is injured. In the following subsections we present the general procedure of Sailfish optimizer algorithm (Shadravan, Naji, & Bardsiri, 2019).

Initialization

- A sailfish and sardine population is initialized randomly, to each sailfish and sardine is assigned a randomized position x_i^k and y_j^k consecutively. Where, i∈ {sailfishes}, j∈ {sardines} and k ∈{number of iteration}
- The position of each sailfish (x_i^k) or sardine (y_j^k) represents a feasible solution for the problem at k^{th} iteration.

Evaluation and the elitism procedure

At each new generated population, an evaluation based on the fitness function (F (.)) of the position of each agent search (sailfish or sardine) is carried out. In a minimization problem, the best sailfish which has the smallest fitness from the sailfish population is saved as the elite sailfish (x_{eli}^k, eli ∊ {set of sailfish}) i.e. F (x_{eli}^k) <= F (x_i^k), ∀ k. Similarly, the best sardine which has the smallest fitness in the sardine population is saved as the injured sardine (F (y_{iinj}^k), inj ∊ {set of sardine}) i.e. F (F (y_{iinj}^k),) <= F (F (y_j^k), ∀ k. The saving of the elite sailfish and the injured sardine is equivalent to an elitism procedure because allows us not to lose the good solutions when we update the position of the search agents.

Position updates by the sailfish

Over course of iteration, each member of group from sailfish population can update his position around the school of the prey. The change of the position in the sailfish's behavior is realized by the alternation of attacks strategies or by the capacity to occupies empty space around the prey school. Mathematically, the update position of sailfish is based principally on the position of elite sailfish and the injured sardine as it's shown in Equation (1)

$$x_i^{k+1} = x_{eli}^k - \lambda_k * (\beta * ((x_{eli}^k + y_{inj}^k)/2) - x_i^k \qquad (1)$$

where x_i^{k+1} is the new position of sailfish at (k+1)th iteration, x_i^k is the current sailfish i position, β is a number generated randomly between 0 and 1, x_{eli}^k andy_{inj}^k are consecutively the position of the current elite sailfish and the position of the current injured sardine and λ_k is a coefficient generated at each k^{th} iteration by the Equation (2)

$$\lambda_k = (2 * \beta * PD) - PD \qquad (2)$$

where, β is a number generated randomly between 0 and 1 and PD presents the density of the school prey. Due to the alternation of attacks on the prey school some sardines will be injured and be hunted by the sailfishes therefore the number of prey will decrease over time. The PD parameter is determined by the Equation (3)

$$PD = 1 - (N_{sh}/(N_s + N_{sh}) \qquad (3)$$

where N_{sh} and N_s are respectively the number of sailfish and the number of sardines in each iteration.

Position updates by the sardine

At the beginning of the hunt, both the sailfish power attack and the sardine escape ability are usually very high. Therefore, at the first stage of the hunt, the sailfishes just injure the sardines in school prey without being able to catch them. Over time, the attack power of the sailfishes decreases as well as the ability to escape from the sardines. Indeed, the sailfishes accuse the effort of the alternation of attacks strategies while the sardines accuse the injuries in their body. This leads to that, at the last stage of the hunt preys to lose the ability to escape from the attacks. Hence, the capture success rate of sailfishes becomes high. To take into account the behavior of sardine against the attacks of sailfish. Each sardine in our algorithm can update its position based on the Equation (4)

$$y_j^{k+1} = r * (x_{eli}^k - y_j^k + AP) \qquad (4)$$

where y_j^{k+1} and y_j^k are the new and the current position of sardine j consecutively, r is a random numbers between 0 and 1, $x_{eli}^k$is the best position of elite sailfish found so far, and AP shows the amount of sailfish's Attack Power at each iteration that is generated as shown in Equation (5)

$$Ap = A * (1 - (2 * Itr * \varepsilon)) \qquad (5)$$

where A and ε are two factors that decrease the value of power attack (AP) and Itr is the number of current iteration. As mentioned previously, at the first stage of the hunt, the capture success rate is small because most of sardines can change their positions and escape from the sailfish attack. However, at the end of the hunting, the ability of sardine to escape decreases which makes the capture success rate higher. So that, the number of sardines that can update their position decreases with time. In our algorithm the last stage of hunting was considered when AP is less than 0.5. Based on power attack AP, the number of sardine that update their position at last stage of hunt (AP < 0.5) is given by Equation (6)

$$\alpha = N_s * AP \qquad (6)$$

where AP is less than 0.5 just a selected number α of sardines can update their position. On the contrary, when AP is more than 0.5 we considered that all sardine will be updated.

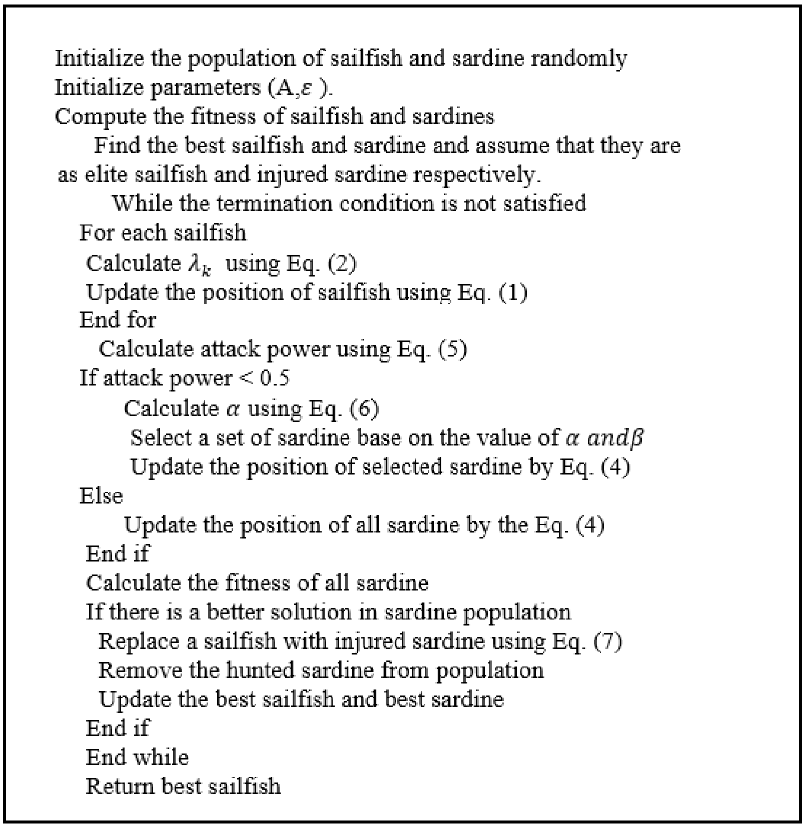
Remove and substitution of the hunted sardine

In the proposed algorithm, to model the fact that a sardine j is hunted by a sailfish i, the position of this later is replaced by the position of the sardine j when we have

$F(y_j^k) < F(x_i^k)$ as given by the Equation (7).

$$x_i^k = y_j^k \text{ if } F(y_j^k) < F(x_i^k) \qquad (7)$$

Where x_i^k is the position of sailfish i at k^{th} iteration and y_j^k is the position of sardine j at k^{th} iteration. After the substitution, the sardine j should be removed from the sardine population. The algorithm of SFO is shown in Figure 3.



```
Initialize the population of sailfish and sardine randomly
Initialize parameters (A,$\varepsilon$ ).
Compute the fitness of sailfish and sardines
     Find the best sailfish and sardine and assume that they are
as elite sailfish and injured sardine respectively.
       While the termination condition is not satisfied
   For each sailfish
    Calculate $\lambda_k$  using Eq. (2)
    Update the position of sailfish using Eq. (1)
   End for
     Calculate attack power using Eq. (5)
   If attack power < 0.5
         Calculate $\alpha$ using Eq. (6)
          Select a set of sardine base on the value of $\alpha\ and\beta$
          Update the position of selected sardine by Eq. (4)
   Else
         Update the position of all sardine by the Eq. (4)
    End if
    Calculate the fitness of all sardine
    If there is a better solution in sardine population
      Replace a sailfish with injured sardine using Eq. (7)
      Remove the hunted sardine from population
      Update the best sailfish and best sardine
    End if
    End while
    Return best sailfish
```

Figure 3. The algorithm of SFO.

Literature review

The main task of steganography is to thwart the unauthorized user from knowing that there is something hidden even he/she can get the stego cover. The section discusses some of the previous work that can help in improving the proposed method by highlighting the points of strength and weakness of it. The proposed method has overcome all the disadvantages that have been shown in the previous methods.

This paper proposes a new method by converting the cover image into gray scale image and then segmentation it, also convert the secret information into sequence of bytes and using LSB (least significant bit) to embedded secret data. This knowledge is represented by the key of image segmentation, the key of mapping within each area segment, data distribution inside segment, key agreement of cryptography method, the key of secret message length and the key of message extension. The main limitation of this technique is limited capacity to embed secret data and imperceptibility (Bawaneh & Obei dat, 2016). For growing the capability of conceal data in RGB image, a method is used to merge the characteristics of color intensities with a Triple-A algorithm where based on randomization. This method is applied on the RGB image where each pixel consists of three colors red, green, and blue. Storage is based on the value of the pixel if it holds smaller values than it would enable to conceal big numbers of data bits. To store data bits, this algorithm uses the real color of the pixel with the generator for pseudorandom. This paper displays good results, essentially in the capability of conceal the data – bits with relation to the pixels of color image. The first drawback of this technique is quality, where changing any bit may effect on the resolution of the image [Rahman, Mondal, Mandal, & Sultana, 2016). For increasing the security, combinations between two techniques are used, where the first technique is RC6 encryption and the second technique is steganography. It also becomes a best method when it uses hybrid encryption techniques. The second drawback of this approach is that it is complex computationally while the experimental dataset is limited (Shrivastava & Singh, 2016). This paper relied on the edges of image for concealment process as human cannot see small changes in this place. It also uses one-time pad algorithm to encrypt the secret data before hiding it wherever the process of hiding is achieved by using least significant bit on the image edge area to increase the imperceptibility. The third drawback of this technique is that there are many applications that can be used to detect the secret message through comparing between cover image and stego image (Irawan, Setiadi, Sari, & Rachmawanto, 2017). This paper proposed a new method to hide text in text by using structure of the invisible character (white space), where the secret messages are hidden in the white space position of the cover image, the results display that the method presents high security because of using high complexity to avoid the unauthorized users. The limitation of this technique is the length of binary string that must be less than or equal to the number of word spaces (Mustafa, 2020). This paper proposed a privacy channel selection rule-based steganography technique in spatial domain. The Just Noticeable Difference (JND) value is merged with a gradient to find a pixel which need to be changed in data hiding process, since JND reveals the minimum visible threshold of HVS. While AG supports to identify the texture characteristic of image region which the pixel belongs to. Therefore, the security of stego-images produced by the new scheme, against statistical steganalysis methods is enhanced. The results show that the stego image of the proposed technique has higher security and quality. The limitation of this technique is the using of the spatial domain, where it becomes vulnerable against the attacks and the lacking of statistical analysis techniques [Nguyen & Le, 2020). A new technique is proposed in this paper, where the secret message (plain text) is first converted into cipher text by using the algorithm of RSA. In the next step, the cipher text is concealed in an audio using the LSB audio steganographic technique.

The proposed schema

Steganography one of the most important technologies used to protect secret information sent via internet, it hides the presence of the secret information instead of encrypting it.

The main idea of the proposed system is to utilize SFO to determine all the accurate matches between the fitness values of the secret message and cover message, also the SFO is considered the key that is used in the embedding and extracting process. This method overlooked the limitations of LSB method in the term of payload capacity and overlooked the complex operation and calculation in the other steganography methods.

The proposed method can be classified into four stages are: calculate the fitness value will be shown in section 'Find the fitness value for secret message and cover message', embedding process will be shown in section 'Embedding process', extracting process will be explained in section 'Extracting process' and evaluation measurements will be demonstrated in section 'Evaluation Criterion'. The flowchart of the proposed method is shown in Figure (4). Both the transmitter and recipient will has the same framework

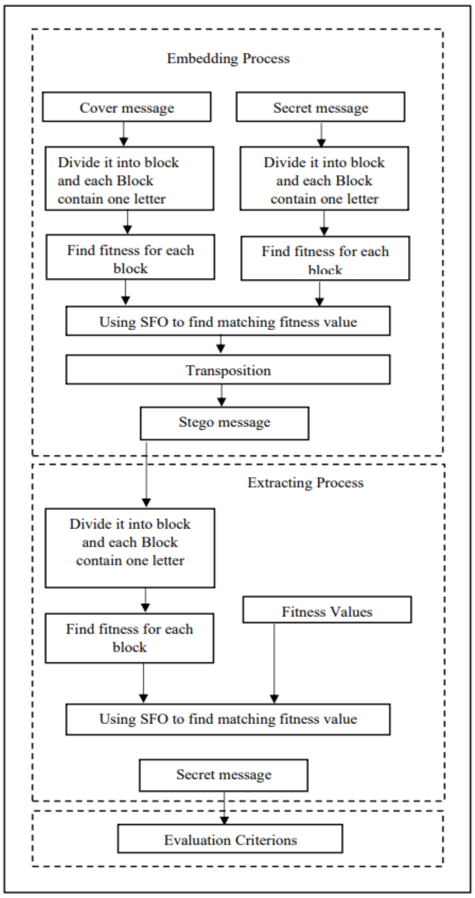

Embedding Process
Cover message
Secret message
Divide it into block and each Block contain one letter
Divide it into block and each Block contain one letter
Find fitness for each block
Find fitness for each block
Using SFO to find matching fitness value
Transposition
Stego message
Extracting Process
Divide it into block and each Block contain one letter
Fitness Values
Find fitness for each block
Using SFO to find matching fitness value
Secret message
Evaluation Criterions

Figure 4. Flowchart of Proposed method.

Find the fitness value for secret message and cover message

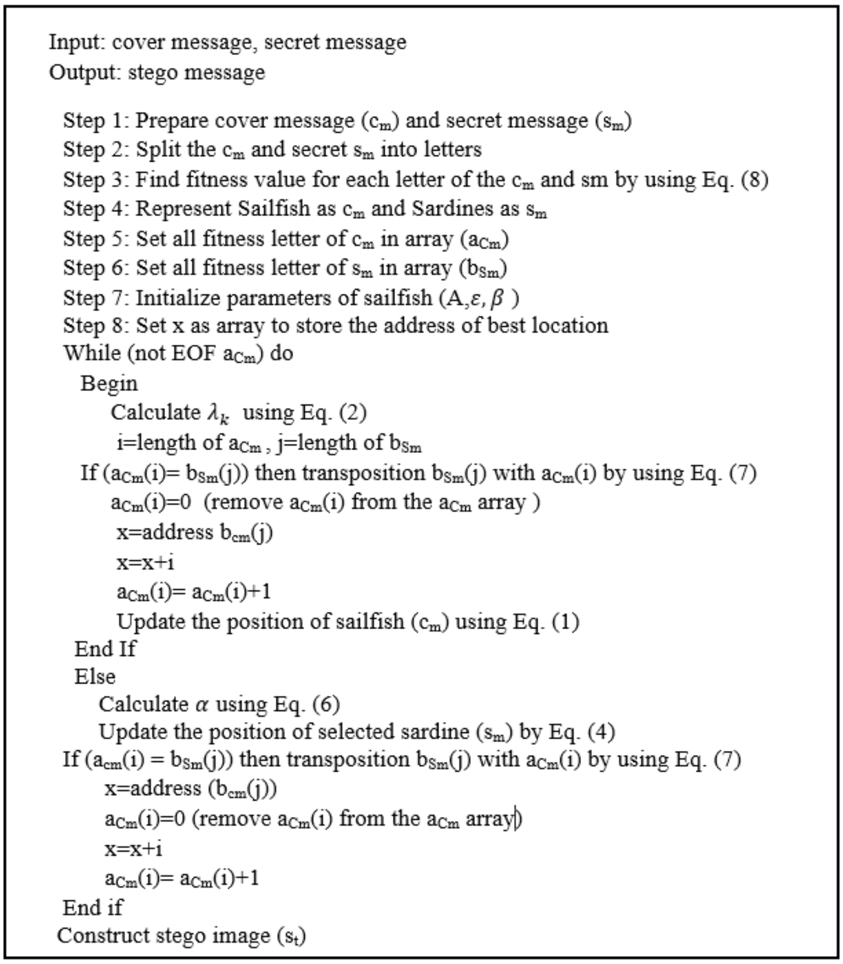
It is one of the important stages in the proposed system where the fitness values for cover and secret message are found, since the SFO based on these values to get the matching between the cover and secret message. This stage begins with divided the secret and cover message into blocks and each block consists of one letter, calculate the sequence of each letter according to the other letters, converting each letter into ASCII code and determine the location of each letter. The fitness value calculated from the following quation:

$$F\ (i)= \frac{A(i)*S(i)}{\sum_i L} * R \qquad (8)$$

Where the A(i) is the ASCII code of the letter, S(i) represent the sequence of letter in the set of letters, L(i) represent the location of letters and the R is random value distribution in the interval [0-1].

Embedding process

At this stage, after find the matching fitness values between the cover and secret messages by using SFO, all matching fitness value and its positions will be saved, the hiding process is done by replacing (transport) secret message letters with some cover message letters. The cover message's letter use one time for transport, if there is more than one letter to transposition, the SFO will decide which one to choose based on its movement, for example if the cover message have 5 letter have the same fitness value, the SFO will decide to choose from them. The key is generated by the positions of the letter in the cover letter which will be used to replace with the secret message. However the key only has one value is its location in the matrix (row, column), therefore the second function of the SFO is considered the key that is used for embedding and extracting process. The embedding process do not effect on the quality of the stego message because there is no operation of add, delete or update. The details of embedding process is summarized in the algorithm (1) Figure 5



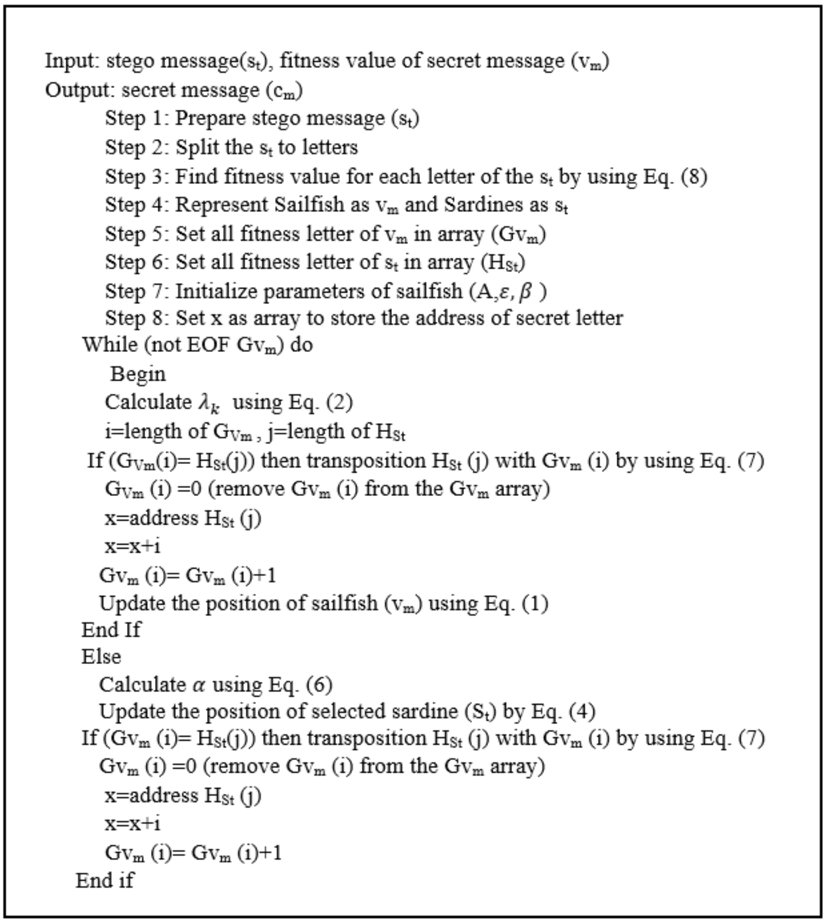
```
Input: cover message, secret message
Output: stego message

  Step 1: Prepare cover message ($c_m$) and secret message ($s_m$)
  Step 2: Split the $c_m$ and secret $s_m$ into letters
  Step 3: Find fitness value for each letter of the $c_m$ and sm by using Eq. (8)
  Step 4: Represent Sailfish as $c_m$ and Sardines as $s_m$
  Step 5: Set all fitness letter of $c_m$ in array ($a_{Cm}$)
  Step 6: Set all fitness letter of $s_m$ in array ($b_{Sm}$)
  Step 7: Initialize parameters of sailfish ($A, \varepsilon, \beta$ )
  Step 8: Set x as array to store the address of best location
  While (not EOF $a_{Cm}$) do
    Begin
        Calculate $\lambda_k$ using Eq. (2)
         i=length of $a_{Cm}$ , j=length of $b_{Sm}$
    If ($a_{Cm}$(i)= $b_{Sm}$(j)) then transposition $b_{Sm}$(j) with $a_{Cm}$(i) by using Eq. (7)
        $a_{Cm}$(i)=0  (remove $a_{Cm}$(i) from the $a_{Cm}$ array )
         x=address $b_{cm}$(j)
         x=x+i
         $a_{Cm}$(i)= $a_{Cm}$(i)+1
         Update the position of sailfish ($c_m$) using Eq. (1)
   End If
   Else
       Calculate $\alpha$ using Eq. (6)
       Update the position of selected sardine ($s_m$) by Eq. (4)
  If ($a_{cm}$(i) = $b_{Sm}$(j)) then transposition $b_{Sm}$(j) with $a_{Cm}$(i) by using Eq. (7)
       x=address ($b_{cm}$(j))
       $a_{Cm}$(i)=0 (remove $a_{Cm}$(i) from the $a_{Cm}$ array)
       x=x+i
       $a_{Cm}$(i)= $a_{Cm}$(i)+1
  End if
 Construct stego image ($s_t$)
```

Figure 5. Algorithm (1) Embedding process.

Extracting process

Extraction process is the inverse of embedding process, at this stage secret message is extracted from stego message by using SFO. The details of extracting process is summarized in the algorithm (2) Figure 6.



```
Input: stego message($s_t$), fitness value of secret message ($v_m$)
Output: secret message ($c_m$)
        Step 1: Prepare stego message ($s_t$)
        Step 2: Split the $s_t$ to letters
        Step 3: Find fitness value for each letter of the $s_t$ by using Eq. (8)
        Step 4: Represent Sailfish as $v_m$ and Sardines as $s_t$
        Step 5: Set all fitness letter of $v_m$ in array ($Gv_m$)
        Step 6: Set all fitness letter of $s_t$ in array ($H_{St}$)
        Step 7: Initialize parameters of sailfish ($A, \varepsilon, \beta$ )
        Step 8: Set x as array to store the address of secret letter
    While (not EOF $Gv_m$) do
         Begin
        Calculate $\lambda_k$  using Eq. (2)
        i=length of $G_{Vm}$ , j=length of $H_{St}$
     If ($G_{Vm}$(i)= $H_{St}$(j)) then transposition $H_{St}$ (j) with $Gv_m$ (i) by using Eq. (7)
        $G_{Vm}$ (i) =0 (remove $Gv_m$ (i) from the $Gv_m$ array)
        x=address $H_{St}$ (j)
        x=x+i
       $Gv_m$ (i)= $Gv_m$ (i)+1
       Update the position of sailfish ($v_m$) using Eq. (1)
    End If
    Else
       Calculate $\alpha$ using Eq. (6)
       Update the position of selected sardine ($S_t$) by Eq. (4)
    If ($Gv_m$ (i)= $H_{St}$(j)) then transposition $H_{St}$ (j) with $Gv_m$ (i) by using Eq. (7)
       $Gv_m$ (i) =0 (remove $Gv_m$ (i) from the $Gv_m$ array)
        x=address $H_{St}$ (j)
        x=x+i
        $Gv_m$ (i)= $Gv_m$ (i)+1
   End if
```

Figure 6. Algorithm (2) Extracting process.

Evaluation criterion

It is impossible to find any difference between the cover message and stego message when comparing them because the hiding process is replaces the secret message's letter with matching (alternative) cover message's letter by using SFO. Therefore, it is impossible for an attacker to know there is hidden secret message or even to suspect it, even if the attacker uses all statistical measures, detection techniques or made a comparison between them visually or through size.

Experimental results and discussion

In this section the results will be presented and discussed, where the results proved the efficiency and effectiveness of the proposed method. The results of the proposed system showed the following:

- The similarity ratio between the stego and cover message is very large.

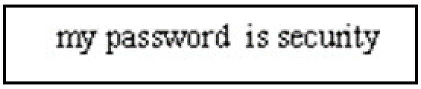
- There is no process of delete, add or update on the cover message

- The hiding process did not effect on the image's quality because the confidential information was hidden with replacement (transposition) method.

- Every letter of secret message has more than one alternative location in the cover message of concealment

- The size of cover and stego message is equal

- The location (matching fitness values) chosen by the SFO are random and chosen smartly. Figure (7) shows the secret message with the size 13.9KB that should be hidden in the cover message, which consists of 4 words and the 4 words consist of 20 letter. The fitness value of the each secret message letter was calculated according to Equation (8), where a new method (previously uninformed) was used in the calculation of fitness values. The Table (1) shows the fitness values of the each secret message letter and also shows that there is no similarity between these values and all values are confined between (0,1)



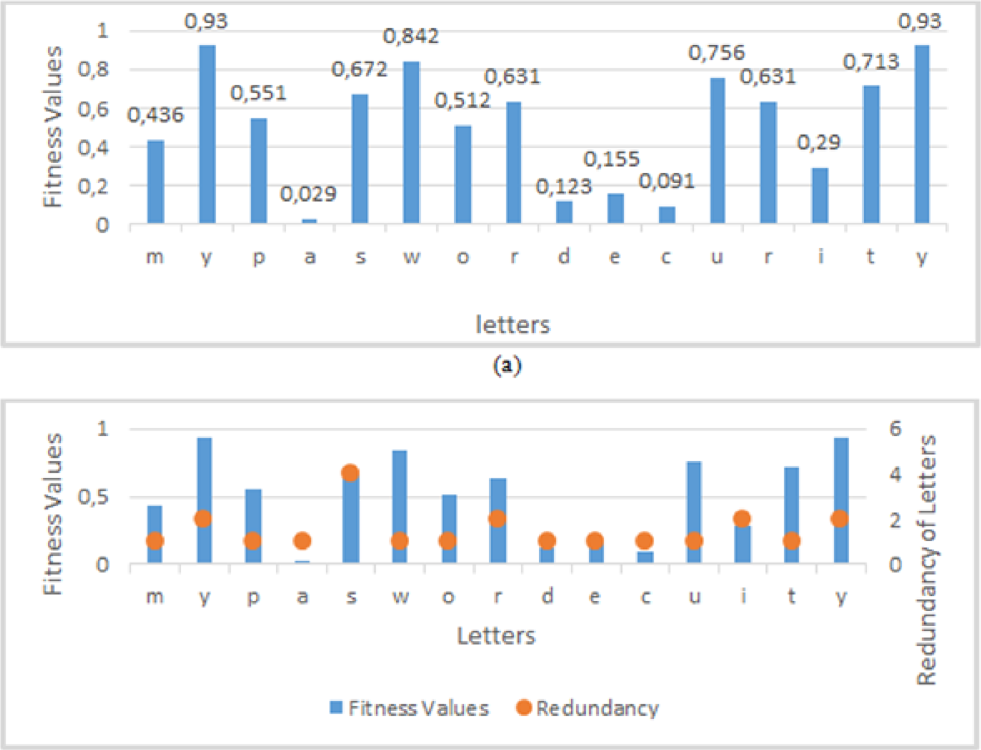
my password is security

Figure 7. Secret message.

Table 1. Fitness value for the secret message.

Sequence	Letters	Fitness Value	Sequence	Letters	Fitness Value
1	m	0.436	11	i	0.290
2	y	0.930	12	s	0.672
3	p	0.551	13	s	0.672
4	a	0.029	14	e	0.155
5	s	0.672	15	c	0.091
6	s	0.672	16	u	0.756
7	w	0.842	17	r	0.631
8	o	0.512	18	i	0.290
9	r	0.631	19	t	0.713
10	d	0.123	20	y	0.930

The relation between letters and its fitness values and between the letters and its redundancy of the secret message is shown in Figure (8), where the letter (y) has the highest fineness value and the letter (a) has the least fitness value also the letter (S) has highest redundancy and letter (a) has least redundancy

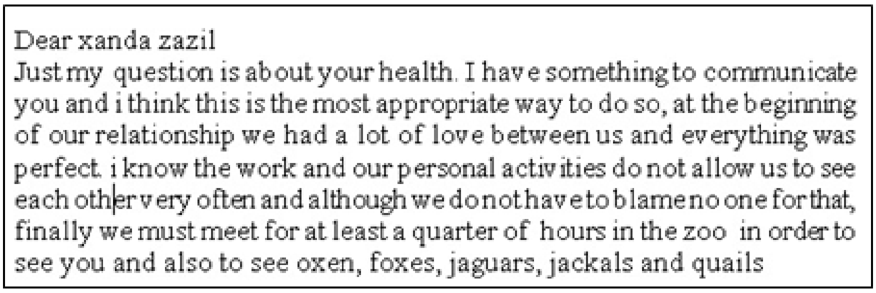


(a)

(b)

Figure 8. Relations of Secret Message: a- Relation between letters and its fitness value, b- Relation between letters and its redundancy.

Figure 9. Shows the cover message with size of 14.4KB, which consists of 105 words and the 105 words consist of 419 letters. The fitness value of the each cover message letters was calculated according to Equation (8) as shown in the Table (2), where each letter of the cover message has fineness value differ from the fitness values of other letters. The same letter that is found in cover message and secret message has the same fitness value, for example the fitness value of the letter (a) in the cover message is the same as the fitness value of the letter (a) in the secret message and is equal to 0.029



Dear xanda zazil
Just my question is about your health. I have something to communicate you and i think this is the most appropriate way to do so, at the beginning of our relationship we had a lot of love between us and everything was perfect. i know the work and our personal activities do not allow us to see each other very often and although we do not have to blame no one for that, finally we must meet for at least a quarter of hours in the zoo in order to see you and also to see oxen, foxes, jaguars, jackals and quails

Figure 9. Cover message.

Table 2. Some of the fitness value for the cover message.

Sequence	Letters	Fitness Value	Sequence	Letters	Fitness Value
1	d	0.123	11	a	0.029
2	e	0.155	12	z	0.976
3	a	0.029	13	i	0.290
4	r	0.631	14	l	0.398
5	x	0.886	15	j	0.326
6	a	0.029	16	u	0.756
7	n	0.473	17	s	0.672
8	d	0.123	18	t	0.713
9	a	0.029	19	m	0.436
10	z	0.976	20	y	0.930

The relation between cover letters and its fitness values and between the letters and its redundancy is shown in Figure (10), where each letter has its own fitness and its redundancy differ from letter to another

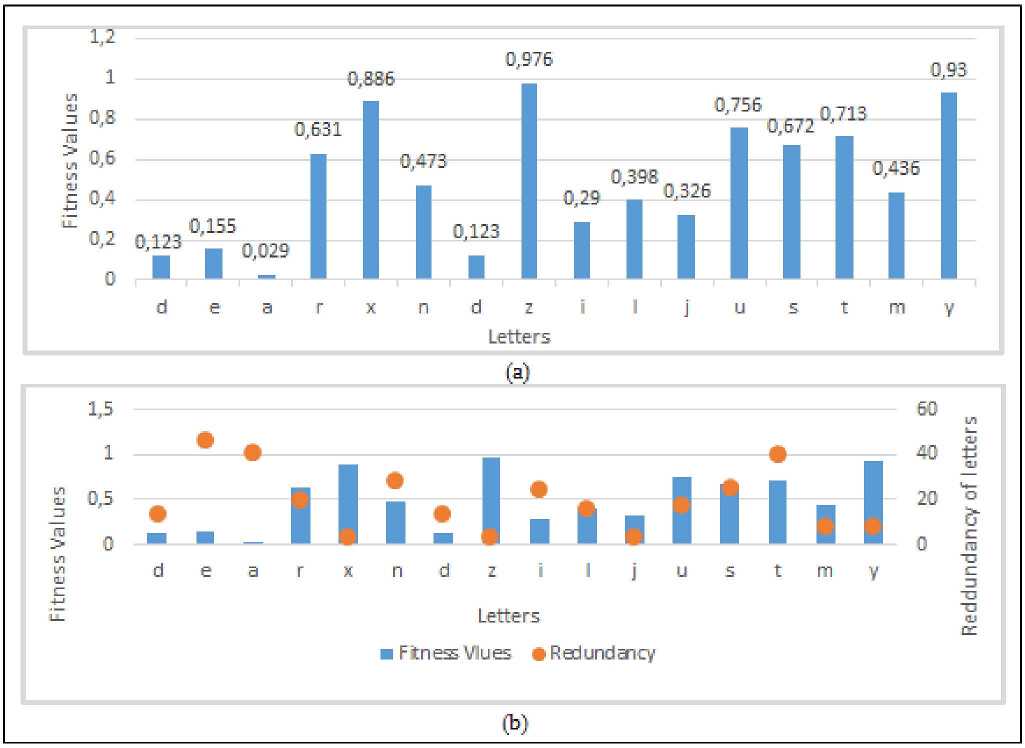


Figure 10. Relations of Cover Message: a- Relation between letters and its fitness value, b- Relation between letters and its redundancy.

Table (3) and Figure (11) shows the number of secret letters and the number of repetitions in the cover and secret message based on the matching fitness values between them, whereas the letter (o) in the secret message has the highest frequency in the cover message and equal to 47 frequency and the letter (c) has the least frequency and equal to 6.

Table 3. Matching fitness values between secret and cover image.

Sequence	Secret letters	Redundancy In secret message	Redundancy In Cover message
1	m	1	8
2	y	2	8
3	p	1	6
4	a	1	41
5	s	4	25
6	w	1	9
7	o	1	47
8	r	2	20
9	d	1	13
10	e	1	46
11	c	1	6
12	u	1	17
13	i	2	24
14	t	1	40

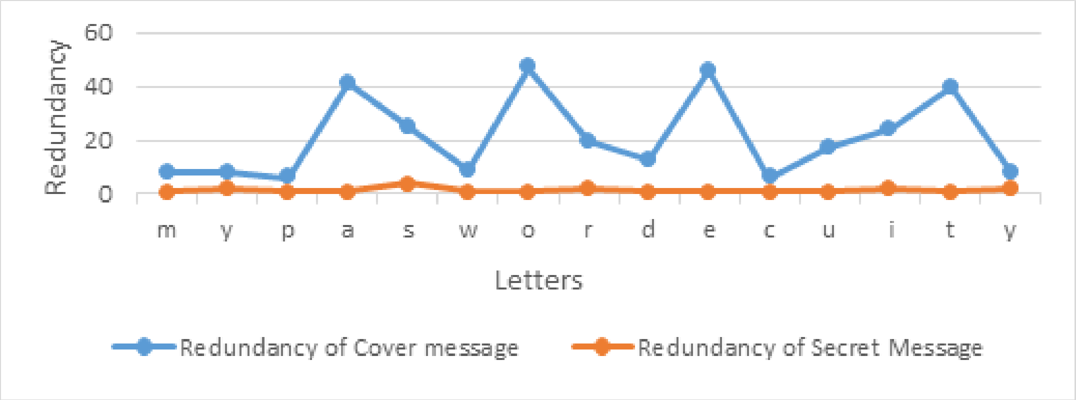
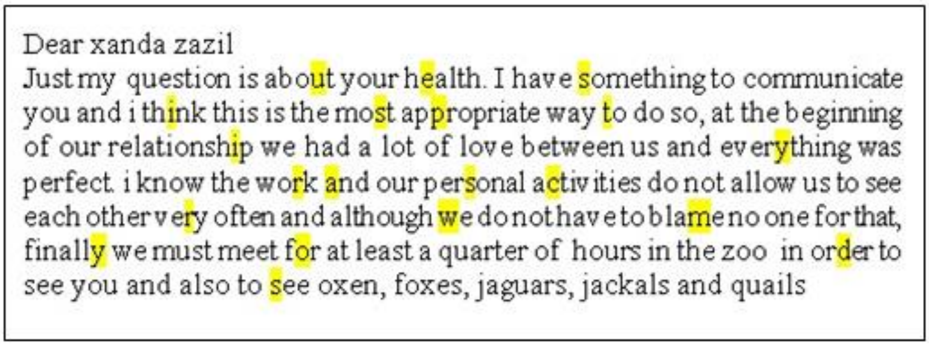


Figure 11. Relation between redundancy of letters in cover and secret message.

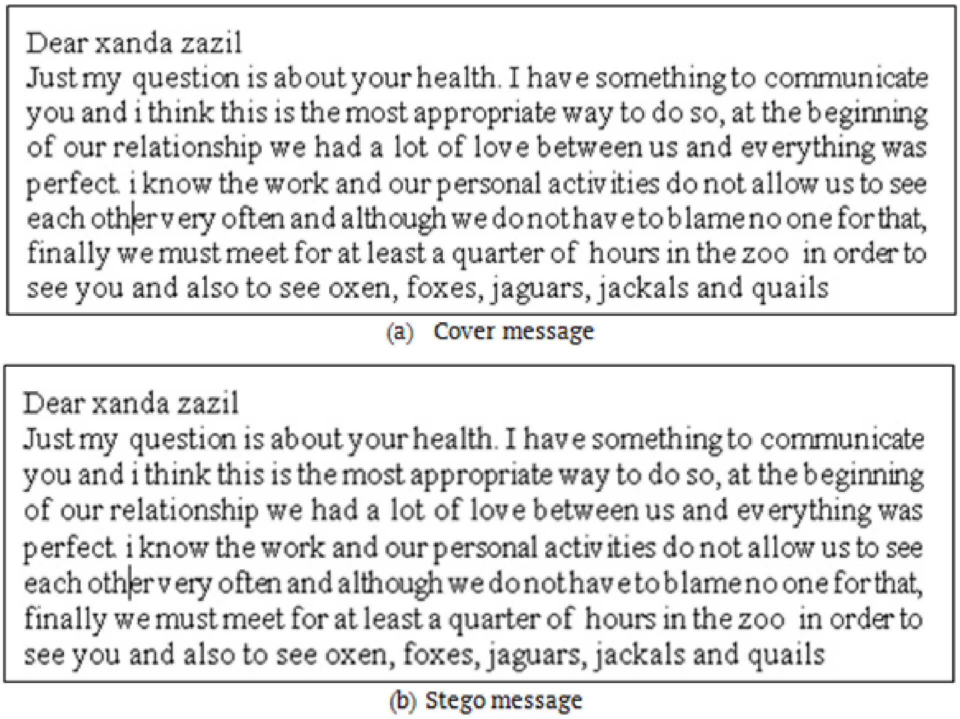
The cover message can accommodate the secret message despite the slight difference between the size of the cover and secret message, where each letter's fitness value of the secret message has several matching (alternatives) letter's fitness value in the cover message. The Figure (12) shows the locations of the cover message's letter (yellow shaded) which was chosen by SFO algorithm to replace (Transposition) with secret message's letter. The similarity between the cover and stego message is very high as shown in Figure (13), that indicates to strength and effectiveness of the proposed method.

Results of Entropy, Energy, Variance, Squared Pearson Correlation Coefficient (SPCC), Structural Similarity Index Metric (SSIM), Signal to Noise Ratio (SNR), Average, Security, Euclidian distance and Chi-square are shown in Table (4), the results of the table shows the strength and efficiency of the method used.



Dear xanda zazil
Just my question is about your health. I have something to communicate you and i think this is the most appropriate way to do so, at the beginning of our relationship we had a lot of love between us and everything was perfect. i know the work and our personal activities do not allow us to see each other very often and although we do not have to blame no one for that, finally we must meet for at least a quarter of hours in the zoo in order to see you and also to see oxen, foxes, jaguars, jackals and quails

Figure 12. The locations of cover letters that used to transport.



Dear xanda zazil
Just my question is about your health. I have something to communicate you and i think this is the most appropriate way to do so, at the beginning of our relationship we had a lot of love between us and everything was perfect. i know the work and our personal activities do not allow us to see each other very often and although we do not have to blame no one for that, finally we must meet for at least a quarter of hours in the zoo in order to see you and also to see oxen, foxes, jaguars, jackals and quails

(a) Cover message

Dear xanda zazil
Just my question is about your health. I have something to communicate you and i think this is the most appropriate way to do so, at the beginning of our relationship we had a lot of love between us and everything was perfect. i know the work and our personal activities do not allow us to see each other very often and although we do not have to blame no one for that, finally we must meet for at least a quarter of hours in the zoo in order to see you and also to see oxen, foxes, jaguars, jackals and quails

(b) Stego message

Figure 13. Similarity between messages: a-cover message, b- stego message.

Table 4. Comparison between the cover message and stego message by using several measurements.

Measurements	Cover message	Secret message	Difference
Entropy	4.12	4.12	0.0
Average	108.7777	180.777	0.0
Energy	2.564	2.564	0.0
Variance	98.58223	98.58223	0.0
SNR	1.9579	1.9579	0.0
Chi-square , p		0, 1	
SSIM		1	
SPCC		1	
Euclidian distance		0	
security		1	

The Table (5) shows the comparison between the proposed method and (Mustafa, 2020), through entropy, average, energy, variance and SNR measurements, through the results of the table, it was found that the proposed method achieved results outnumbered the results of the mentioned paper.

Table 5. Comparison between proposed system and paper (Mustafa, 2020).

Measurements	(Mustafa, 2020)	Proposed method
Entropy	6.4838E-00006	0.0
Average	1.4344E-00002	0.0
Energy	2.5068E-00007	0.0
Variance	0.0	0.0
SNR	-1.2045E-0000	0.0
Payload capacity	Increase payload capacity	Increase payload capacity
Text quality	Not decrease	Not decrease
Text hidden	Low text hidden	More text hidden
Time	More consume	Low consume
Secure	Less	high

The Table (6) shows the comparison between the proposed method and other methods through accuracy of security, through the results, it was found that the proposed method achieved results outnumbered the results of the mentioned paper.

Table 6. Comparison of security between methods and the proposed system.

Approach	Security
(Suhad, 2016)	0.90
(Abdulraheem, 2014)	0.93
(Abdulwahed, 2020)	0.99
(Mustafa, 2020)	0.997
(Nguyen & Le, 2020)	0.999
Proposed method	1

Conclusion and future work

Hiding text message in text message by using sailfish algorithm is a novel and efficient approach and provides high security. Experimental results show the following features of the proposed approach:

-Each letter in the cover and secret message has a fitness value unlike any other fitness value and the same letter that found in the secret and cover message at the same time has the same fitness values as shown in table (1) and table (2)

-Provide high capacity, where each letter of the secret message matching several letters in the cover message as shown in Table (3).

-Provide high security, as the SFO chose the concealment locations (matching fitness values) with intelligent and randomly way (not sequence) as shown in Figure (12).

-There is no difference between cover and stego message because the concealment process is only transpose between them as shown in Figure (13).

-The size of cover message is the same size of stego message (no change in the size) so it does not need special format or font.

-Resists most attack methods as shown in Table (4)

-The proposed method achieved impressive results through use set of metrics and also performed on the other papers as shown in Table (5) and Table (6).

-To detect the secret message, the attacker needs to know the algorithm use, number of iteration, parameters value and the initialize location of sailfish algorithm.

-The stego quality remained good during the transmission and not affected with noise and corruption as shown in Figure (13).

-A proposed method is achieved by lower-case English alphabets. It can be also expanded to work for many language as turkey, Japanese etc., upper case, symbols and numbers.

References

Abdulwahed, M. N.(2020). An effective and secure digital image steganography scheme using two random function and chaotic map. *Journal of Theoretical and Applied Information Technology, 98*(1), 78-91.

Alwan, Z. A., Farhan, H. M., & Mahdi, S. Q. (2020). Color image steganography in YCbCr space. *International Journal of Electrical and Computer Engineering (IJECE), 10*(1), 202-209. DOI: http://doi.org/10.11591/ijece.v10i1.pp202-209

Baawi, S. S., Mokhtar, M. R., & Sulaiman, R.(2017). New text steganography technique based on a set of two-letter words. *Journal of Theoretical and Applied Information Technology, 95*(22), 6247-6255.

Bawaneh, M. J., & Obei dat, A. A. (2016). A secure robust gray scale image steganography using image segmentation. *Journal of Information Security, 7*, 152-164. DOI: http://doi.org/10.4236/jis.2016.73011

DOI: http://doi.org/10.9790/0661-0341117

Hamdan, A. M., & Hamarsheh, A. (2016). AH4S: an algorithm of text in text steganography using the structure of omega network. *Security and Communication Networks, 9*, 6004–6016. DOI: http://doi.org/10.1002/sec.1752

Irawan, C., & Setiadi, D. R. I. M., Sari, C. A., & Rachmawanto, E. H. (2017). Hiding and securing message on edge areas of image using LSB steganography and OTP encryption. In *Proceedings of the 2017 1st International Conference on Informatics and Computational Sciences* [ICICoS], (p. 1-6). DOI: http://doi.org/10.1109/ICICOS.2017.8276328

Kaur, H., & Rani, J. (2016). A Survey on different techniques of steganography. *MATEC Web of Conferences 57*, 02003. DOI: http://doi.org/10.1051/matecconf/20165702003.ICAET

Kingslin, S., & Kavitha, N. (2015). Evaluative approach towards text steganographic techniques. *Indian Journal of Science Technology, 8*(29), 1–8. DOI: http://doi.org/10.17485/ijst/2015/v8i1/84415

Maleki, N., Jalali, M., & Jahan, M. V. (2014). Adaptive and non-adaptive data hiding methods for grayscale images based on modulus function. *Egyptian Informatics Journal, 15*(2), 115–127. DOI: http://doi.org/10.1016/j.eij.2014.06.001

Mohamed, A. A. (2014). An improved algorithm for information hiding based on features of Arabic text: a Unicode Approach. *Egyptian Informatics Journal, 15*, 1-9. DOI: http://doi.org/10.1016/j.eij.2014.04.002

Mustafa, N. A. A. (2020). Text hiding in text using invisible character. *International Journal of Electrical and Computer Engineering (IJECE)*, 10(4), 3550~3557. DOI: http://doi.org/10.11591/ijece.v10i4.pp3550-3557

Nguyen, T. D., & Le, H. D. (2020). A secure image steganography based on JND model. *International Journal of Electrical and Computer Engineering (IJECE), 10*(2), 2088-2096. DOI: http://doi.org/10.11591/ijece.v10i2.pp2088-2096

Rahman, M. M., Mondal, P. K., Mandal, I., & Sultana, H. (2016). Secure RGB image steganography based on triple-A algorithm and pixel intensity. *International Journal of Scientific & Engineering Research, 7*(3), 864-869.

Roy, S., & Venkateswaran, P. (2013). A text based steganography technique with indian root. In *Proceedings of the 2013 International Conference on Computational Intelligence: Modeling Techniques and Applications* [CIMTA], (p. 167-171). *Procedia Technology, 10*, 167– 171. DOI: http://doi.org/10.1016/j.protcy.2013.12.349

Setyono, A., & Setiadi, D. R. I. M. (2019). Securing and hiding secret message in image using XOR transposition encryption and LSB method. *Journal of Physics: Conference Series, 1196*, 012039. DOI: http://doi.org/10.1088/1742-6596/1196/1/012039

Setyono, A., Setiadi, D. R. I. M., & Muljono, M. (2017). StegoCrypt method using wavelet transform and one-time pad for secret image delivery. In *Proceedings of the 2017 4th International Conference on Information Technology, Computer, and Electrical Engineering* [ICITACEE], (p. 203-207). DOI: http://doi.org/10.1109/ICITACEE.2017.8257703

Shadravan, S., Naji, H. R., & Bardsiri, V. K. (2019). The sailfish optimizer: a novel nature-inspired metaheuristic algorithm for solving constrained engineering optimization problems. *Engineering Applications of Artificial Intelligence, 80*, 20-34. DOI: http://doi.org/10.1016/j.engappai.2019.01.001

Shrivastava, A, & Singh, L. (2016). A new hybrid encryption and stenography technique: a survey. *International Journal of Advanced Technology and Engineering Exploration, 3*(14), 9-14. DOI: http://doi.org/10.19101/IJATEE.2016.314005

Shukur, W. A., & Jabbar, K. K. (2018). Information hiding using LSB technique based on developed PSO algorithm. *International Journal of Electrical and Computer Engineering (IJECE), 8*(2), 1156-1168. DOI: http://doi.org/10.11591/ijece.v8i2.pp1156-1168

Singh, P., Chaudhary, R., & Agarwal, A. (2012). A novel approach of text steganography based on null spaces. *IOSR Journal of Computer Engineering (IOSRJCE), 3*(4), 11–17.

Suhad, M. K. (2016). Text steganography method based on modified run length encoding. *Iraqi Journal of Science*, *57*(3C), 2338-2347.

Taha, M. S., Rahim, M. S. M., Lafta, S. A., Hashim, M. M., & Alzuabidi, H. M. (2019). Combination of steganography and cryptography: a short survey. *IOP Conference Series: Materials Science And Engineering, 518*(5). IOP Publishing.