Scientific Journal of Impact Factor(SJIF): 3.134

e-ISSN(O): 2348-4470
p-ISSN(P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 2,Issue 5, May -2015

Tree-based Data Hiding Technique Using Genetic Algorithm

Sharon Sunny
Dept. of CSE, AmalJyothi College f Engineering, sharonsunny@amaljyothi.ac.in

Abstract — *An important concern in 'covered writing' or steganography is to reduce the distortion of cover image caused by embedding secret message. At the same time, it should preserve the integrity of data. A simple and efficient technique that can be adopted for achieving this goal is tree based parity check(TBPC) method. An advancement of this scheme is majority-vote parity check(MPC). In this paper, I propose a refinement of MPC method. Instead of simply embedding the secret text in cover image, after dividing the message and cover image into equal number of blocks, we will see which block of message can be best mapped to which block of cover image. Stego image can be obtained by combining the stego blocks.*

Keywords-: *steganography, genetic algorithm, tree-based steganography, majority vote parity check, toggling*

I. INTRODUCTION

Data hiding techniques have been in use for many years. Military applications find extensive use of this technique. But world came to know about the power, the destructive power of this technique after 9/11 attack. Do not have a mis-interpretation that technology is dangerous.

Steganography is the technique of hiding data in a digital media with such a perfection that a third person could not have any hell idea of a hidden message. The basic terminologies include cover object, secret message and stego object. Cover object refers to the medium in which data can be hidden. Text, image, audio files or video files [4] are examples of cover object. In this work color image is being used as a medium, which has an added advantage over grey scale images that it provides three layers for embedding data, in effect increasing the capacity of the medium. Once data is hidden into cover object, it can be called as stego object. The technique of steganograpy include two phases: an embedding phase and extraction phase. Embedding algorithm explains how secret message can be hidden in cover object where as extraction algorithm explains how the same can be extracted from stego object [11]. Care should be taken while embedding data since changes to the cover object should be minimum. The proposed method focuses on this criteria.

A classic steganographic model, the prisoner's problem [5],[6] is presented by Simmons. Alice and Bob are two prisoners planning to escape together. All communications between them are sent through public channel and is monitored by the jail warden, Wendy. So messages must be hidden within some innocuous-looking media. Wendy can actively or passively monitor the communication. If she is a passive warden, she will inspect messages to determine the presence of any hidden message. If she is active, she will modify the hidden message, if any detected.

A wide variety of techniques have been proposed earlier for data hiding. Wet paper codes [3] embed the data in random pixel positions which may make the decryption difficult even for the intended recipient. In double layered method for data embedding [7] proposed by Weiming Zhang et al., a pixel can carry secret bits by choosing adding/subtracting one to/from the gray value. Bin Li et al. described various methods of image steganography like Least Significant Bit (LSB) based steganography, multiple bit-planes based steganography, and prediction error based steganography [1]. LSB based steganographic techniques make changes in the LSB of pixel values to hide secret data. Pixels may be chosen randomly.This technique is used in our proposed method. Tools [2] like Steghide, S-tools, Steganos, etc are available for implementing LSB based steganography. Multiple bit-planes based steganographic methods alter not only LSB but other bit planes also. This technique increases the capacity of embedding but may reduce the quality of stego image. Prediction error based steganography [1] hide the secret data in complex areas of image. Pixel prediction error can be used to find the complex areas.

A digital image may be considered as a matrix of pixel values. A pixel can take values from 0 to 255. Pixel value can be represented using 8 bits. LSB contain no visual information and a human eye cannot detect any change in its value [9],[4]. Steganographic techniques that uses digital images as cover objects take advantage of this weakness. R. Y. M. Li et al. proposed an LSB steganography scheme based on tree structure called tree-based parity check (TBPC) [8]; to reduce distortion on a cover object.

Block based methods can be adopted as a betterment of LSB based methods. Instead of simply embedding the secret data in cover image, message blocks are embedded after checking resemblance with cover image blocks. Let N denote the number of blocks. Mapping function between cover image and secret message blocks can be a complete permutation of N. When the value of N is too small, the problem can be solved in polynomial time. N should be large enough to get full advantage of blocking scheme; but then the problem may become non-polynomial time. In this paper, a solution to this problem is sought with the help of GA.

1.1. Genetic algorithm

While travelling around the world in HMS Beagle, Charles Darwin noticed biodiversity in Galapagos Island. This inspired him to come up with The Theory of Evolution. Genetic algorithm (GA) is a computational model inspired by this theory. Creation of initial population is first step in the implementation of GA. Each member in population will be a string called as chromosome, which encode the potential solution to a particular problem [13]. Initial population is generated randomly. After generating the initial population, each chromosome is checked with a fitness function. Only the fittest chromosomes will survive, and be chosen for the creation of next generation. Execution of GA is a two stage process. Intermediate population is generated by applying a selection procedure to current population. Next population is generated by applying recombination and mutation over intermediate population.

II. Existing Method

In TBPC method, an N-ary tree called master tree is constructed from LSB of pixel values. The nodes are filled from left to right and from top to bottom. Let L be the number of leaves in master tree. A master string of length L can be obtained from this tree by performing parity check on each path from root to leaves (e.g. see Fig.1). Then secret message bits are XOR-ed with master string. Resultant toggle string is placed in leaf nodes of a toggle tree. Internal nodes of this toggle tree are filled with zeros. Then for each level, from bottom to top, a non-leaf node and its child nodes are flipped if all the children have bit 1. Stego tree can be obtained by XORing master tree and final toggle tree. Extraction algorithm extract the secret data from stego tree by simply performing parity check on all the paths of from root to leaves. In Majority-vote Parity Check (MPC) [10] the toggling criteria of nodes is relaxed. A non-leaf node and its child nodes are flipped if majority of its children have bit 1; thereby reducing the number of 1s in toggle tree. Number of 1s in toggle tree denotes the number of bit changes for embedding the secret message bits. Hence lesser the number of 1s in toggle tree, lesser will be the modification of cover object.

II. EXISTING SYSTEM

In TBPC method, an N-ary tree called master tree is constructed from LSB of pixel values. The nodes are filled from left to right and from top to bottom. Let L be the number of leaves in master tree. A master string of length L can be obtained from this tree by performing parity check on each path from root to leaves (e.g. see Fig.1). Then secret message bits are XOR-ed with master string. Resultant toggle string is placed in leaf nodes of a toggle tree. Internal nodes of this toggle tree are filled with zeros. Then for each level, from bottom to top, a non-leaf node and its child nodes are flipped if all the children have bit 1. Stego tree can be obtained by XORing master tree and final toggle tree. Extraction algorithm extract the secret data from stego tree by simply performing parity check on all the paths of from root to leaves. In Majority-vote Parity Check(MPC) [10] the toggling criteria of nodes is relaxed. A non-leaf node and its child nodes are flipped if majority of its children have bit 1; thereby reducing the number of 1s in toggle tree. Number of 1s in toggle tree denotes the number of bit changes for embedding the secret message bits. Hence lesser the number of 1s in toggle tree, lesser will be the modification of cover object.

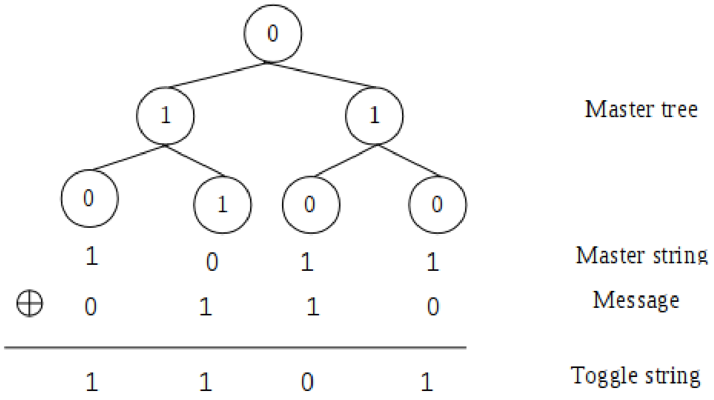
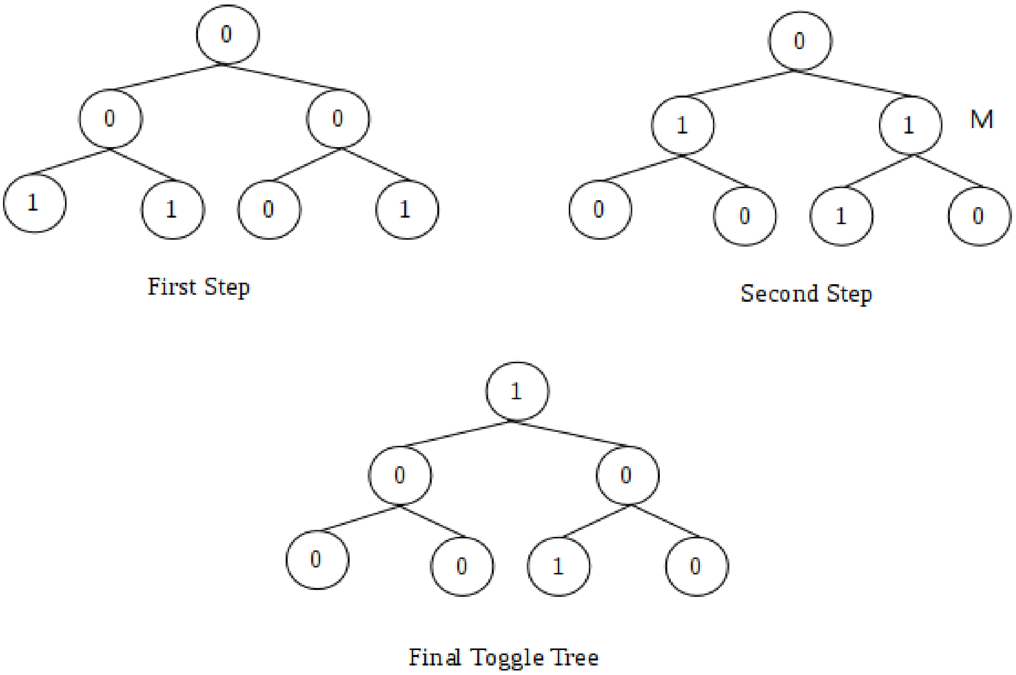


Fig.1 Master and toggle string of a master tree with L=4 for LSBs 0,1,1,0,1,0,0 of the cover image

III. PROPOSED METHOD

Two critical issues in a steganographic system are: 1. reducing the number of modifications of cover object thereby reducing distortion, and 2. better efficiency for embedding and extraction [10]. Our method tries to reduce distortion on cover object by finding a mapping function between cover image block and message block.

To begin with, divide the message and cover image into equal number of blocks, say N. There would be N! ways possible for mapping. Obviously evaluation of all these permutations for finding best mapping function may not be completed in polynomial time. We can utilise GA to solve this problem.

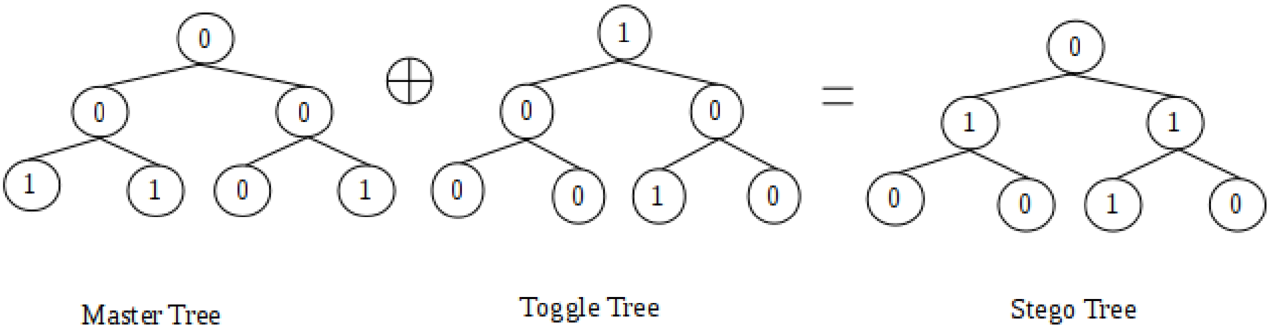


First Step

Second Step

Final Toggle Tree

Fig.2 Construction of a toggle tree with L=4 for toggle string 1,1,0,1



Master Tree

Toggle Tree

Stego Tree

Fig.3 stego tree constructed from the toggle string 1, 1, 0,1

Embedding algorithm may be described as follows:

```
1. Divide cover image and secret message into N blocks ( as shown in Fig.4).
2. Generate M initial chromosomes each of size N. The value inside each sequence is produced randomly and has no repeats.
4. For i= 1 to M   // for each chromosome
        4.1. Choose ith chromosome as mapping function.
        4.2. For j= 1 to N // for each snippet in chromose. Image blocks are taken in the order
                  specified in chromosome
                4.2.1. Construct master tree using the LSB of pixel values of cover image block (e.g.  See Fig. 1). Nodes are filled from left to right and from top to bottom. Generate master string from the tree. Secret message bits and master string are XORed to generate toggle string.
                4.2.2. Construct toggle tree whose leaf nodes are occupied by toggle string and internal nodes by 0s (e.g. see Fig.2).  Toggling criteria of MPC can be adopted for generating final toggle tree.
                4.2.3. Construct stego tree by XORing master tree and toggle tree (e.g. see Fig. 3).
        4.3. Combine N stego image blocks.
        4.4 Find PSNR value for the stego image generated
5. Out of M chromosomes, K chromosomes with highest Peak Signal to Noise Ratio ( PSNR) values are chosen for reproduction. Crossover and mutation are applied to get the next population. Repeat 4 L times
6. From the final population, chromosome that generate stego image with highest PSNR is chosen and is sent to receiver.
```

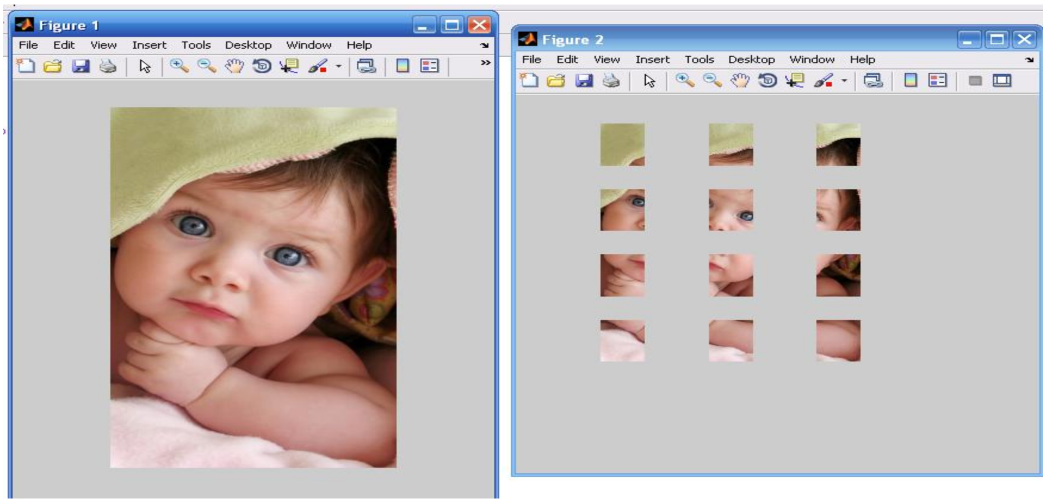


Figure 1
File Edit View Insert Tools Desktop Window Help
Figure 2
File Edit View Insert Tools Desktop Window Help

Fig. 4 Cover image and blocks of image

Let us explain the algorithm with the help of an example. Let image and message are divided into say 4 blocks. Let the mapping functions chosen randomly are [4 2 1 3], [2 1 3 4], [1 2 3 4] . The above algorithm will take the first mapping function [4 2 1 3]. This mapping function means 4th block of message will be mapped to 1st block of image, 2nd message block will be mapped to 2nd image block and so on. The same procedure is repeated with other chromosomes. We choose say 2 chromosomes for the genneration of next population. Selection is soley based on PSNR value. Next population may yield a better chromosome thereby optimizing the problem.

For extracting the secret message receiver, should have the mapping function with him. The following procedure may be followed while extracting the message.

```
1. Divide stego image received into N blocks.
2. For i = 1 to N          // for each stego image block
        a. Construct master tree as explained above.
        b. Check the parity on each path from root ro leaves to get the message block embedded in ith block.
3. Combine message blocks in the order specified in mapping function to get the original message.
```

This method can be adopted when secret message is of comparatively larger size. Use of GA will significantly reduce distortion of cover image since it helps to find an optimal mapping function.

IV. ANALYSIS

Root-mean-square (RMS) error, MSE and peak signal to-noise ratio (SNR) PSNR are two customary criterions to evaluate quality of a processed image. They are described as

$$M_{SE}=\frac{1}{MN}\sum_{1\leq i\leq M}\sum_{1\leq j\leq N}\left(f'(i,j)-f(i,j)\right) \qquad (1)$$

where M is the number of rows in the image matrix and N is the number of columns in image matrix.

$$PSNR=10\log_{10}\frac{255^2}{M_{SE}} \qquad (2)$$

In this paper, PSNR is chosen for evaluating the quality of stego image. Higher PSNR value indicates that stego image is close to cover image. Table below shows the comparison of proposed method with MPC method and GA based LSB substitution method.

Table 1. PSNR comparison

GA based LSB substitution	MPC	TBPC	Our method
45.7826	46.4018	46.0123	46.6714

A 1080*1024 image was used for simulation. The obtained PSNR values are shown in Table 1. System is implemented using Matlab. Time complexity of the proposed method can be expressed in polynomial time.

V. CONCLUSION

In this paper, I propose a novel method for embedding secret message in a cover object. Proposed method is a blend of tree based parity check method, blocking method and LSB substitution method. Inorder to get an optimal solution genetic algorithm is also incorporated. Performance of this method was compared with existing methods and the result was satisfactory.

Drawback of this scheme is that performance of GA is affected by the selection of initial population. Also adaptive block dividing strategy may yield a better result.

REFERENCES

[1] Bin Li, Junhui He, and Jiwu Huang, “A Survey on Image Steganography and Steganalysis”, Journal of Information Hiding and Multimedia Signal Processing, Volume 2, Number 2, April 2011.

[2] http://www.stegoarchive.com

[3] Jessica Fridrich, Miroslav Goljan, Petr Lisoněk, and David Soukal,”Writing on Wet Paper”,

[4] Ali Daneshkhah, Hassan Aghaeinia, and Seyed Hamed Seyedi, “A More Secure Steganography Method in Spatial Domain”, 2011 Second International Conference on Intelligent Systems, Modelling and Simulation

[5] G. J. Simmons, “Prisoners’ problem and the subliminal channel,” in Proc. Int. Conf. Advances in Cryptology, 1984, pp. 51–67.

[6] Yi-Ta Wu and Frank Y. Shih,”Genetic Algorithm Based Methodology for Breaking the Steganalytic Systems”, IEEE Transactions on Systems, Man, And Cybernetics—Part B: Cybernetics, Vol. 36, No. 1, February 2006

[7] Weiming Zhang, Xinpeng Zhang, and Shuozhong Wang “A double layered plus-minus one data embedding scheme”, *IEEE Signal Processing Letters*, Volume 14, No. 11, November 2007.

[8] R. Y. M. Li, O. C. Au, K. K. Lai, C. K. Yuk, and S.-Y. Lam, “Data hiding with tree based parity check”, *Proc. IEEE Int. Conf. Multimedia and Expo (ICME 07)*, pp. 635–638, 2007.

[9] Rafael C. González, Richard Eugene Woods, “Digital Image Processing”, Third edition, Pearson Education Inc 2008.

[10] Chung-Li Hou, ChangChun Lu, Shi-Chun Tsai, and Wen-Guey Tzeng, “An Optimal Data Hiding Scheme With Tree-Based Parity Check ”, IEEE Transactions On Image Processing, Vol. 20, No. 3, March 2011

[11] Jayesh George.M, j.jayageetha, “An optimal data hiding scheme with block-based parity check for binary, gray and colour images”, IJART, Vol.2 Issue 5,May 2012, 11- 19.

[12] Rongrong Ji, Hongxun Yao, Shaohui Liu, Liang Wang, “Genetic Algorithm Based Optimal Block Mapping Method for LSB Substitution”, Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing.