Covert Transfer of Biometric Data using Steganography

Srinidhi G A* and K B ShivaKumar**
*Research Scholar, Sri Siddhartha Academy of Higher Education, Tumakuru, Karnataka, India
srinidhiga@ssit.edu.in
** Research Supervisor, Sri Siddhartha Academy of Higher Education, Tumakuru, Karnataka, India
shivakumarkb@ssit.edu.in

Abstract: "The secret between three people will remain secret provided, any two are dead" is a proverb which was coined by Benjamin Franklin. Today, the secured transfer of information is a very basic need. The rapid use of interconnected computing systems gives way for many different types of eavesdropping and has become a profession for many hackers. So there is a need for the data being trans received to be having a secured mode of communication. Earlier, these authentication was required just for delegates of higher order in the society as they were supposed to carry their transactions in a secured mode. But today, the security of data is of need even by a common lay man. For example, the issue of sim cards was taking place after the through verification of person's identity details as a hard copy. Whereas today, every transactions are digital in nature and it has reduced the time taken to authenticate the person so as to activate the service which he/she has requested for. But on the other end, the issue of security is of major concern. This paper proposes a new technique using both cryptography and steganography methods which makes the covert data much secured. The results obtained in the simulation phase are so promising and are better than the ones existing in the current literature.

Keywords: Steganography, Covert Communication, Crytography..

Introduction

The word steganography has a history dates back to the ages of kings' rule. Then, many physical commodities were used to secure the data being transmitted. One such method was to use milk as an invisible ink to write and form patterns on the paper so that it would not be visible for the normal human eye. The trick here is to expose the same in front of the fire or any other light so as to see the characters or the patterns visible. The other method of doing it was using a secret messenger. A secret messenger was a normal human being whose head was shaved so as to tattoo the secret information to be sent. He was kept in prison for some days so that his hairs will grow back. This person was sent to the other end of communication so that he was shaved back to get the secret data. The same technique has been employed here in a digital form which contain a cover media and a secret media. Both of them are in digital form which makes it possible to hide one inside the other.

A typical example of such is using a digital image as both cover media and also the secret data which is referred as payload.

As said, both of them are in digital form. The pay load is considered which nothing but the pixel values to be considered. Here the major information about the color of the pixel is concentrated in the higher nibble of the pixel value and these four bits of the pixel is secretly transmitted, the payload image at the receiver end may be fairly reconstructed.

In the cover image, the data will again be in the form of pixels and is similar to the payload image. Further, if the lower nibble of the cover image is altered in a logical way, the stego data looks alike the cover image so that the very purpose of steganography is achieved.

Steganography can be considered as an advanced version of cryptography. Cryptography is nothing but scrambling of data so that the data becomes unreadable for most of the users. But there are two major problem with this encryption scheme. The first one is that there is no chance that the data is completely secured as the data embedded may be decrypted by using every possible trails so that the data can be decrypted. Anyhow the time constraint is a major factor which makes cryptography still in use even today.

The second drawback of crypto system is that, the existence of the covert data is not concealed. Whereas this drawback has been overcome by the any stego system. Therefore, steganography plays a very important role in the process of securing the covert data being transmitted over a non-secured channel like internet.

It may also be noted that the combination of both cryptography and steganography makes it possible to have much more secured data transfer over the internet.

Figures 1 and 2 shows a typical block diagrams of a simple stego system and the system which has both crypto and stego system together to have a higher security. The blocks here are as explained below:

Cover Data: This can be considered as a dummy data which is used just to carry the covert data in it through the un secured channel.

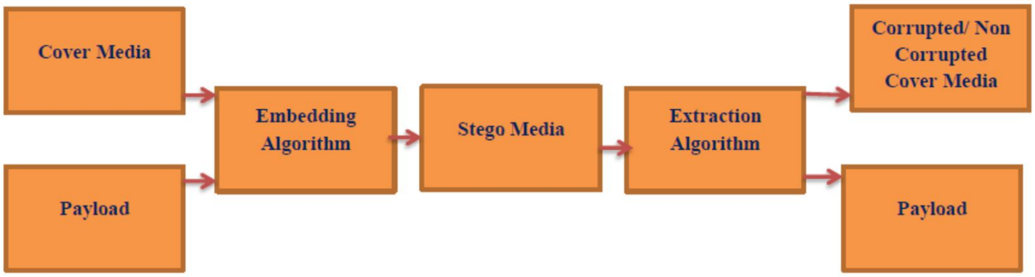


Figure1. Typical Stego System

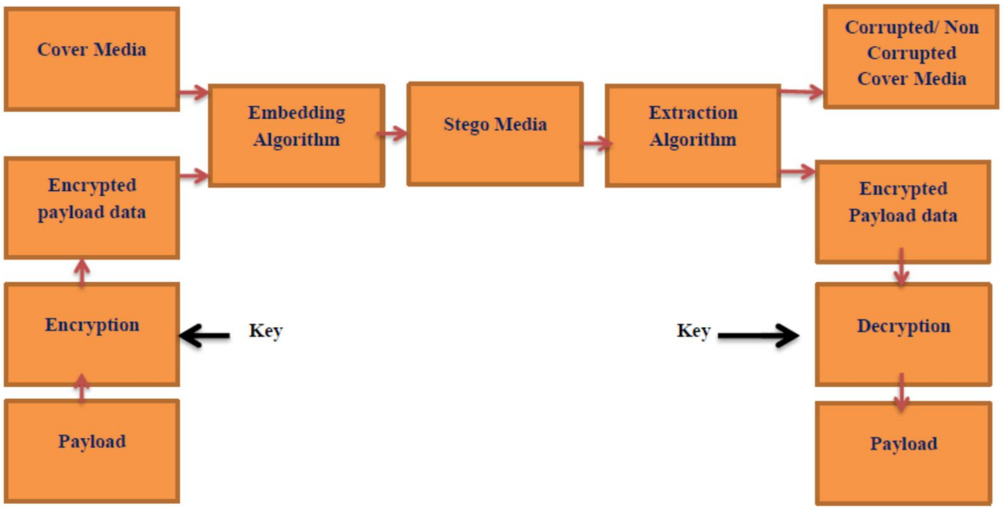


Figure2. Stego System with Crypto System for additional Security

Payload: Payload or covert data is nothing but the secret data being transmitted.
Embedding Algorithm: Embedding algorithm is any algorithm designed to embed the payload data being transmitted into the cover media so that the cover media will not have any noticeable change.
Stego Media: This is the product which is obtained by the investment of cover media, payload and the embedding algorithm. This is done at the transmitter end usually.
Extraction Algorithm: The algorithm which is just the reverse process of embedding algorithm. The output of which is the corrupted or non-corrupted cover media. But fairly good payload data which should be exactly same as the payload which is embedded into the cover at the transmitter end.
Encryption Algorithm: This is the block which is present only in the system which has crypto system in it. Here one or the other cryptography algorithm which is used to encrypt the covert data into a un understandable data.
Decryption Algorithm: This is an algorithm which is the reverse process of encryption algorithm.
Key: A Key is a text which is nothing but the secret data known only for both transmitter and the receiver which is an essential input for the decryption algorithm
Further, The paper reviews the available literature in section 2. Section 3 proposes the algorithm developed to embed the biometric data of human being (Iris image is considered). The results of experimenting the same has been presented in section 3 followed by concluding remarks and brief idea about the future enhancements in section 4.

Literatures Reviewed

Debnath Bhattacharyya et al.,[1] proposed a combined approach for steganography where the text is first encrypted using a key and then it is hidden behind a cover image. Even though many steganography techniques elaborates the same concept in several different way, the paper presents a good explanation of dependency of text and their ASCII as well as UTF values on the overall steganography and also introduces layer wise model.
Indradeep Banerjee et al., [2] presented text steganography method where encryption of the text is followed by hiding it in another text where characters are generated in an order such that they appear random to intruders. However authors have not

emphasized on presenting a meaningful text as outcome of steganography which triggers suspicion. This technique could be more comprehensive if the text is hidden in cover text in such a way that resultant text is also correct text format.

Por and Delina [3] proposed a solution for the problem explained in [2] where text is hidden behind cover text such that result presents some meaning and are not just random sequences. However their statistical based approach needs huge cover text for hiding any text such that randomness of the characters is retained.

Shraddha et al., [4] came up with technique to hide the text message by dividing the letters into two groups of curvatures and non curvature letters. Further the text is hidden inside another text by encoding group ID followed by number of the letter in the group. This technique is quite unique. However it does not clearly elaborate the scenario as how curvatures are naturally identified by their algorithm. Also the technique depends upon the believe that the secret message is atleast a paragraph long. Dependency of the technique on length of the message and the paragraphs limits the technique.

Yousuf Bassil et al., [9] elaborated a technique to hide SQL queries by developing a meta dictionary to map different queries with different random selectively chosen table names and fields. Therefore this provides a good framework for generic text steganography where a method can be developed to append texts suitably into secret message such that not only the secret message can be retrieved at the receiver but at the same time schema of the text or domain of meaning of the text is retained.

Changder et al., [5] presented a novel steganographic technique to hide text using binary domain extraction. They first obtain the binary equivalent of the characters and hence sentences and then a similar sentence of character is obtained from an image that is closer to secret message. The technique still needs to supply a carrier image which makes the Steganography as more hybrid than pure text steganography. However their work on binary domain emphasizes on applying steganography on binary data rather than the text itself. This finding lays the foundation for our work where we use UTF data of the text to hide information. Indradip Banerjee et.al.,[6] proposed a procedure of text steganography using Indian Regional Language. Text steganography together with quantum approach based on the use of two specific characters and two special characters like invited comas (opening and closing) in Oriya language and mapping technique of quantum gate truth table was used to generate the stego text with minimum degradation. In this method the length of stego and cover remain same and this property enables the method to avoid the steganalysis also.

Christine K. Mulunda [7] proposed a genetic algorithm based model in Text Steganography worked with text as the cover medium with the aim of increasing robustness and capacity of hidden data. Elitism is used for the fitness function. The model presented is applied on text files, though the idea can also be used on other file types. M.Grace Vennice et.al.,[8] proposed a technique for hiding the Text Information using Stegnography using inter-word and inter paragraph spacing for hiding information. Shraddha Dulera et. al.,[4] proposed a method based on combining the random character sequence and feature coding methods to hide a character and evaluated the approaches based on metrics viz. hiding strength, time overhead and memory overhead entailed. it.

Proposed System

The system proposed here uses both cryptography and steganography system together. A term hybrid steganography has been coined for this system which makes it possible to have a much better security levels.

Here the system considers the iris image which is a most desirable biometric data and being used in most of the cases after figure print. It should be noted here that the challenge of using iris as a payload data is that, usually t he finger print data is just the gray scale image in most of the cases. But the iris image is an RGB image. Therefore the capacity required by the iris image to get embedded into the cover image almost triples. Also the next challenge with the iris image is that it consists of a larger data which is in terms of MBs and is used. As the proposed application is in authentication, there is no chance that even a single pixel can go wrong at the receiver end. Therefore, the algorithm has been designed to overcome all these challenges and to provide a fair security system for the biometric data being transmitted.

Figure 3 shows a brief flow of the algorithm employed in this work. Here there is a need to explain how exactly the encryption is happening and also how exactly the embedding process is taking place.

Encryption: As the payload data is an RGB image, each pixel is a combination of 3 values each ranging from 0-255. The three components are Red, Blue and Green components. Here, very simple technique is used which divides the red component by 2n Blue component by 2n+1 and the blue component by 2n-1 where n is any natural number ranging between 0-50 which is also a key being used by users at both transmitting and the receiving ends. Here the trick used is that the data being transmitted is RGB and requires a lot of capacity so the data is compressed without even losing it.

Embedding: The embedding process takes place in spatial domain without actually using any of the standard transformation processes. Here the embedding uses LSB steganography but the change is that each pixel used variable number of bits instead of just a single bit. The process here is to consider the difference between the two neighbors. The difference between the pixel values are considered to decide the number of bits to be embedded into the pixel. If the difference is >50, 2 bits are added in each color component of the pixel so that it embeds 6 bits per pixel. If the difference is >100, then 3 bits of covert

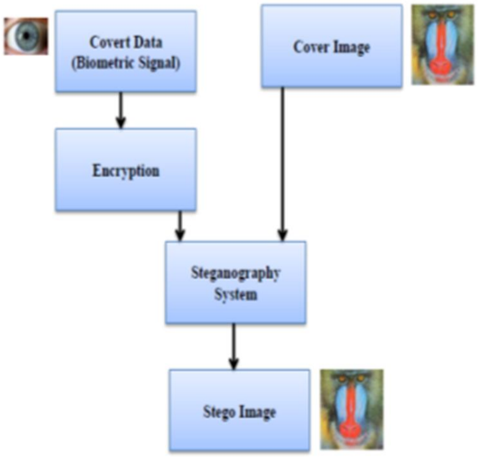


Figure3. Proposed Embedding Algorithm

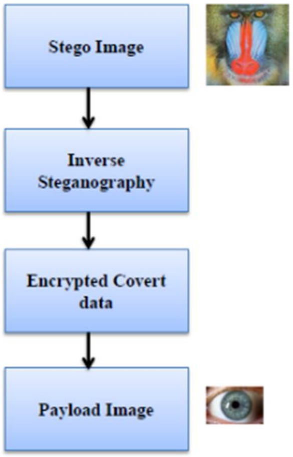


Figure4. Proposed Extraction Algorith

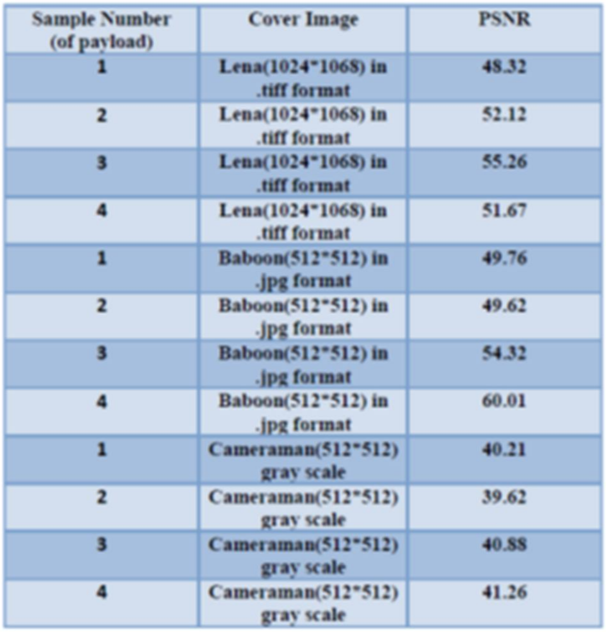
data is embedded into each color component of the pixel so that 9 bits are added in each pixel. If the difference is <50, then just one bit is added to each color component which is same as LSB steganography.

The process of decryption and extraction is exactly the reverse process of the encryption and embedding process and the algorithm is coded to be exactly reverse.

Results and Discussion

The proposed algorithm is simulated using MATLAB 2014 image processing tool box. Images which are most commonly used in almost every steganography algorithms like lenna, camera man, Baboon, pepper etc are used as cover image samples of which has been presented in table 1. The iris image has been downloaded from google image database and has been tested for almost 200 images. Four samples with different combinations of cover image has been presented in table 1. It may be noted that the is always a trade of between the capacity of the cover image and the PSNR which is calculated using MSE using the standard formulae to calculate both of them and are available in the literature. It may also be noted that the algorithm gives scope to use different images of any resolution and any y format can be used as cover image and also iris image of any size and any format can be used as the payload image. Any how a number ranging between 1-50 can be used as the key.

Table 1. Obtained results for different combinations of Covers and Payloads



Sample Number (of payload)	Cover Image	PSNR
1	Lena(1024*1068) in .tiff format	48.32
2	Lena(1024*1068) in .tiff format	52.12
3	Lena(1024*1068) in .tiff format	55.26
4	Lena(1024*1068) in .tiff format	51.67
1	Baboon(512*512) in .jpg format	49.76
2	Baboon(512*512) in .jpg format	49.62
3	Baboon(512*512) in .jpg format	54.32
4	Baboon(512*512) in .jpg format	60.01
1	Cameraman(512*512) gray scale	40.21
2	Cameraman(512*512) gray scale	39.62
3	Cameraman(512*512) gray scale	40.88
4	Cameraman(512*512) gray scale	41.26

Conclusions and Future work

As it is evident from the results as shown in table 1, the highest PSNR obtained is for the combination Baboon of resolution 512*512 which is in .jpg format and the iris sample 2. It may be noted that as per the literature, PSNR above 40 is considered to be the working steganography system.[4,5,8]. But the system gives better PSNR which makes the system promising. The algorithm has also been tested with a gray scale image cameraman where the capacity of it will be less than the regular RGB image also it may be noted that, even though the cover image is gray scale, the RGB payload is embedded. Where in there is a small change in the algorithm where the bits are embedded into the gray level pixels instead of color components. Even then, the results are quite promising and are above the ones being discussed in the literature.

In future, the algorithm is planned to be implemented for real time data where in the iris will be directly scanned from the human eye using a camera. This requires the process of segmentation as well as a preprocessing step before encryption.

References

[1] Debnath Bhattacharyya, Poulami Das, Samir Kumar Bandyopadhyay, Tai-hoon Kim, “Text Steganography: A Novel Approach,” International Journal of Advanced Science and Technology,Vol. 3, February, 2009.

[2] Indradip Banerjee, Souvik Bhattacharyya, Prof. Gautam Sanyal, “Novel Text Steganography through Special Code Generation,” International Conference on Systemics, Cybernetics and Informatics, pp 208-303, 2011.

[3] L. Y. Por, B. Delina, “Information Hiding: A New Approach In Text Steganography,” International Conference on Applied Computer & Applied Computational Science, Hangzhou, China, 2008.

[4] Shraddha Dulera, Devesh Jinwala and Aroop Dasgupta, “Experimenting With The Novel Approaches In Text Steganography,” International Journal Of Network Security & Its Applications, Vol.3, No.6, November 2011.

[5] S.Changder, D. Ghosh, N. C. Debnath, “LCS based Text Steganography through Indian Languages,” International Conference on Computer Science and Information Technology, pp. 53-57, 2010.

[6] Indradip Banerjee, Souvik Bhattacharyya and Gautam Sanyal, “A Procedure of Text Steganography Using Indian Regional Language,” International Journal for Computer Network and Information Security, pp.65-73, 2012.

[7] Christine K. Mulunda, Peter W. Wagacha and Alfayo O. Adede, “Genetic Algorithm Based Model in Text Steganography,” The African Journal of Information Systems, Vol 5, pp. 131-144, 2013.

[8] M.Grace Vennice, Prof.TV.Rao, M.Swapna, and Prof.J.Sasi kiran, “Hiding the Text Information using Stegnography,” International Journal of Engineering Research and Applications, Vol. 2, Issue 1, pp.126-131, 2012.

[9] Youssef Bassil, “A Generation-based Text Steganography Method using SQL Queries,” International Journal of Computer Applications, Volume 57, No.12, November 2012