Volume 4, Issue 9, September 2014 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Data Security with Image Clustering using Hopping Neighbour Technique

Mitali Garg
Computer Science Department
Chandigarh University, Gharuan, India

Vikas Wasson
Computer Science Department
Chandigarh University, Gharuan, India

Abstract*— *In this modern era, internet is widely used for the communication of information. People are easily able to exchange the information but privacy or secrecy of communicated message is an important issue. Steganography or data hiding is the process of hiding the information in such a way so that its existence is not known. The ultimate goal of this thesis is to develop the technique that provides security to the information by hiding it in the image file format using hopping neighbour technique. Here the jpeg image file format is used for hiding the information. In the first step Black and white clusters are formed individually by counting the black and white pixels of the image. The three primary colors i.e. Red, Green and blue are also used for forming the clusters. The largest cluster among all of them is chosen for the hiding purpose. The reason for choosing the largest cluster is that it will not degrade the quality of image. Further hopping neighbour technique is applied for hiding the information in the largest selected cluster. Results show that current technique using hopping neighbour approach provides improved values of hiding capacity, as well as minimized values of time consumption and linear visual distortion as compared to traditional neighbouring approach.

Keywords*— *Steganography, Data Hiding, Cover Object, Cryptography, Steganalysis.

I. INTRODUCTION

In the modern era, computers and the internet are major communication media through which the world is connected. As a result, people are easily able to exchange the information and distance is no longer a barrier to the communication. The safety and security of long-distance communication remains an issue. Data transfer and sharing is a part of high speed Internet technology. Intruders try to access the secret Information. So Information Security is needed to be applied and modified exponentially. Cryptography and Steganography are used for information security. The word Steganography is arrived from Greek word *Steganos* which means "covered" and *Graphie* means "writing". Steganography is the art in which information is hidden in such a way that existence of message is undetectable, so it is also called covert communication [2]. This technique is used to hide secret information and to prevent any attackers to use the Information in illegal form. Here the secret information is embedded in some carrier without leaving any evidence of data alternation and the secret information can be plain text or cipher text. Data Hiding is another name for this technique. The Goal of Steganography is that data should be hidden in such a way that if even viewed by user should not gain focus from the viewer. Cryptography is the classic method of securing the communication. It is used to encrypt data i.e. it generates cipher text from the plain text and make it unreachable for unauthorized person. Cipher text can be easily transmitted but it can arouse suspicion for the attacker which can lead to attack or violent decryption of the message [3]. So now a day's cryptography is most widely used with Steganography. Rise of internet is one of the major developments in last few years and people have easy access to information through the computers. Digital libraries, latest news and information about all over the world are easily available online. Drawback of this is that digital information can be misused by making illegal copies. Hiding of digital information has risen now a day's especially for audio and video files but Steganography or data hiding was used throughout the history and is not a new mean.

Herodotus, the father of history was the person who reported about use of Steganography. He mentioned that wax tablets were used for hiding the text when Sparta's was notified by demeratus that the king of Persia intends to invade Greece and the message could be recovered by scrapping wax off the tablet [6].Some chemical substance was also used for concealing the information and was very common mean of hiding data in history. Organic substances like urine or milk were used to make invisible ink and the secret information could be written with this invisible ink in between the lines and information becomes visible if the document is heated gently. In ancient china, paper masks were carried by people to agree for the locations of the letters having secret information. Mask was placed over the sheet of paper by the sender and then secret information was written in holes and the location was filled up by the cover text and the receiver extracts the information by placing his mask over the message. This technique of sending information was used during World War II by German spy [6]. Steganography and cryptography serves the same purpose of securing the communication but they differ from each other. Three basic criteria's for measuring the steganograpic system are:

Capacity - Maximum length of the secret information is known as capacity .Embedding function and properties of the cover are the factors on which the capacity depends. Absolute terms like bits can be used for specification for the given cover or relative to the number of bits required for storing. All the messages are not of maximum length so the measure used for capacity is bits per pixel. Another metric used is *P* for proportion and is used when the length of secret information is relative to maximum length of cover. Embedding rate *P* is not measured in units, it is defined in the range of $0 \leq p \leq 1$. Different compression formats have different capacity measures [4].

Imperceptibility - The aim of Steganography is basically to hide the existence of communicated message. Unlike cryptography it does not depends on difficulty in reading the content of communicated secret message. Steganography basically depends on undetectability of communicated secret information so that it does not arouse the suspicion of the attacker and it covert the communication. So while embedding data it should be seen that the difference between stego image and original image should be minimal such that the unauthorized person cannot detect Secret information. If more information is hidden inside the carrier image, it will lead to degradation of stego image. So the quality of the image should be checked consistently before embedding large amount of data so that it should not become noticeable. This property is satisfied only when cover object and the stego image are not distinguishable. There are various evaluation techniques but peak signal noise ratio (PSNR) is mostly used [4].

Robustness - Stego image should remain unchanged even if it undergoes transformation like sharpening, filtering, scaling, blurring, cropping and other modification etc. It should produce the original image after reversal of processes. This makes sure that the message to be hidden remains safe even if it gets attacked and manipulated. It is basically how to resist the attacks against distortion which are done intentionally [4]. Steganalysis attacks are also used to evaluate the robustness. The aim of Steganalysis is to extract the secret information and there are numerous methods available.

II. RELATED WORK

Akhtar N.,Khan S. et al. in paper entitled "**An Improved Inverted LSB Image Steganography**" [22] an improvement in the plain LSB based image Steganography is proposed. The use of bit inversion technique is used to improve the stego image quality. The size of cover image is 8 times more for message image which increases the bandwidth to transmit the image. Information can easily be extracted by collecting the LSBs if suspicion of the attacker is aroused for the secret information.

Altaay A.,Sahib S. et al. in paper entitled "**An Introduction to Image Steganography Techniques**" [15] intends to give an overview of image Steganography, its uses and techniques. Author says that information hiding technology falls into three classes of Steganography, watermarking and cryptography. Capacity, robustness and imperceptibility are the three important metrics for evaluating the data hiding techniques. Apart from these three metrics compression ratio, multiple watermarks, success rate, complexity of embedding data and detection complexity are also considered as important metrics for data integrity.

Ashwin S.,Kumar S. et al. in paper entitled "**Novel and Secure Encoding and Hiding Techniques using Image Steganography: A Survey**" [20] describes a short survey on the various techniques of Steganography. It describes that Steganography is the art of passing information in such a way that it does not arouse any suspicion of the transmitted message. Earlier cryptography was used for securing communication but it only encrypts the plain text to generate cipher text which can lead to suspicion so Steganography was used as a mean of covert communication. Author says that Steganography is better mean of securing information than cryptography. Biswas D., Biswas S. et al. in paper entitled "**Digital Image Steganography using Dithering Technique**" [23] proposes an innovative technique for hiding and retrieving the secret image. Here author's main focus is on the digital image Steganography i.e. where main focus is using image as medium to hide the secret information. The main reason behind using images as a cover medium is the higher degree of redundancy present in representation of the digital image. The proposed innovative technique consists of two main processes i.e. encoding and decoding. While decoding it is tried to match the quality of secret image to the original image but still there is loss in the quality of image.

Chandramouli R., Memon S. in paper entitled "**Analysis of LSB Based Image Steganography Techniques**" [8] presents rigorous approach at arriving at the Steganography capacity of LSB based image data hiding techniques. Author says that Steganography is statistical according to the classical definition. It derived expression of probability of false detection for detection purpose in the terms of number of hidden bits. Capacity expression used here is just an upper bound and there are many tests to detect the presence of secret information. Here the author's main focus is on images but the results are equally applicable on multimedia data like audio and video.

Chanu Y., Tuithung T. et al. in paper entitled "**A Short Survey on Image Steganography and Steganalysis Techniques**" [25] presents a survey on various Steganography and Steganalysis techniques. Author describes the various techniques used for Steganography of the images which are in spatial and transform domain and also discusses about the Steganalysis. Here author refers Steganography as an art of covert communication and the goal is to ensure the secret communication is not detectable.BMP file format was also not considerable because of large size as compared with other formats which was not good for transmission purpose.

Cao H.,Kot A. in paper entitled "**On Establishing Edge Adaptive Grid for Bilevel Image Data Hiding**" [24] experimentally proved an efficient method for edge adaptive data hiding. This is method for authenticating binary host images through establishing a dense edge-adaptive grid along the objective contours. It is efficient than IB4 scheme for images with high resolution clarity and good exemption quality. This technique of data hiding provides 77% times improvement in payload capacity..

Cheddad A., Condell J. et al. in paper entitled "**Digital image Steganography: Survey and analysis of current methods**" [5] provides review and analysis of various techniques of Steganography with some guidelines and standards. The word Steganography is originated from Greek language which means "covered writing". Peak signal to noise ratio is used for measuring performance of image distortion and the value below 30 db indicates that the quality is low. These methods have low payload than the spatial domain methods. Familiar images should not be used as cover image and it should be preferred to create own image for this purpose.

III. METHODOLOGY

In this research project of mine the text is embedded in the image file so that information can be easily communicated without arousing any suspicion. The following are the proposed steps for hiding information using hopping neighbour technique.

Step 1: Generate Training Set Samples.

- Input the type of data and image to be secured.

Step 2: Preprocessing of Image to count number of pixels to form Clusters.

- Divide the image into the regions or the parts that are conducive.
- Calculate the grey level of pixels of image.
- Construct the clusters/Regions with pixels having identical values.
- To decide upon which cluster is to be selected for textual data hiding.

Step 3: Securing the data using Hopping Neighbour Method.

- Finding the exact location by calculating the total number of pixels in the cluster and the amount of bytes required to hide the textual data keeping the size in mind.
- Hiding all the data in largest cluster of the 2D image.

Step 4: Extraction process.

- Reveal password
- Read from the stego-image the information that is stored in the pixels.

Step 5: Output the Secret Information.

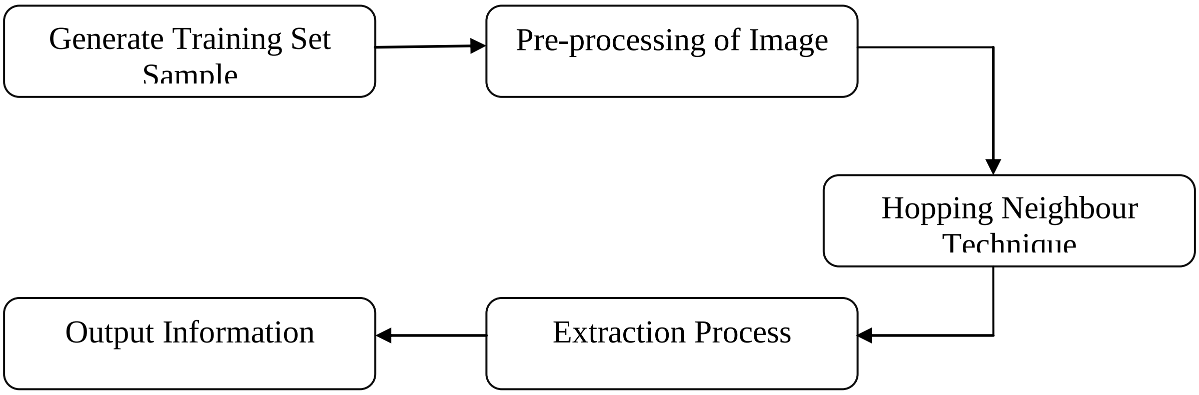


Figure1: Methodology of Data Security

In this process of hiding information, firstly the information to be secured is input along with the image. Then the counting of pixels is done and clusters are formed i.e RGB-WB. Among all, largest cluster is selected for the purpose of hiding the data. The largest cluster is used for the hiding and this is done with the hopping neighbour technique.

	Algorithm: Hiding Process
	Begin *1. Input the .jpeg Image and Information to be stored* *2. Preprocessing of the Image* *3. Construct the Clusters RGB-WB* *4. Select the Largest Cluster* *5. Hopping Neighbour is applied to largest cluster* *End*

Hopping neighbour technique is used for the process of hiding. Password is used for the security purpose. Here largest cluster is selected from all the clusters and then hiding is done. Nearest neighbours are selected within the cluster for hiding the data. For revealing the information password is re-entered, if it matches the initial password the information is revealed otherwise message is displayed.

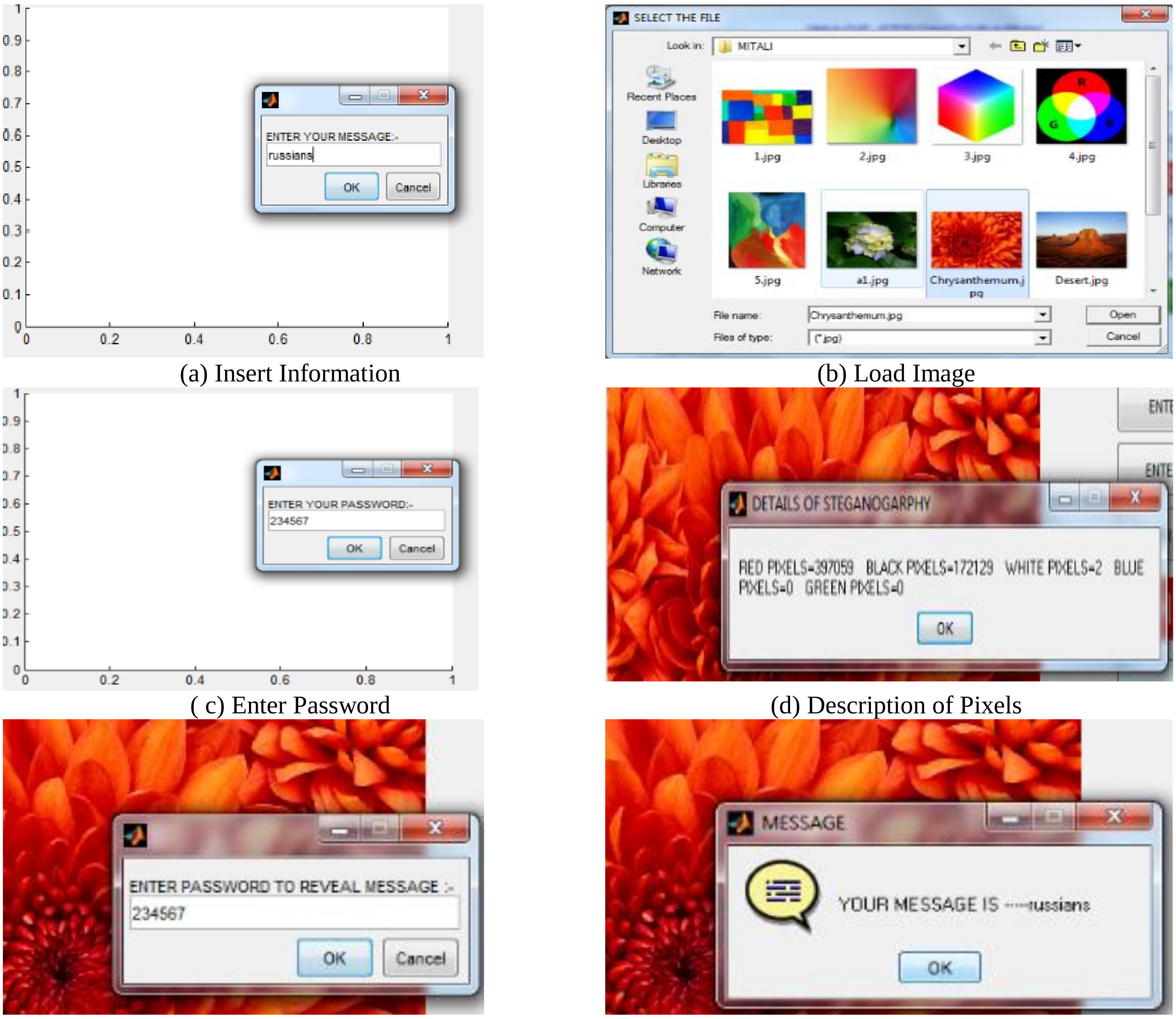
	Algorithm: Hopping Neighbour Technique
	Begin *1. Input the .jpeg Image and Information to be stored* *2. Input password for security p1* *3. Count Pixels of RGB-WB* *4. Cmax= Max RGB-WB* *5. Count= Size of Cmax*

```
6. Select the nearest neighbours for hiding
7. Reveal password p2
8. If {
   P1=P2
   Reveal Information
 9. Else
 Wrong password entered
}
End
```

IV. RESULTS AND COMPARISONS

The current section presents the results obtained after implementing Hopping Neighbour Technique on images. The screen shot results of the different steps of Hiding Process are presented. The comparison results Hopping Neighbour Technique and Neighbouring Technique are presented in Table. Finally for quick view of the results and better understanding all these comparison results are presented in form of Graphs. All the implementation is done using Matlab.

Screen Shot Results



ENTER YOUR MESSAGE:-

russians

File name: Chrysanthemum.jpg

Files of type: (*.jpg)

(a) Insert Information (b) Load Image

ENTER YOUR PASSWORD:-

234567

DETAILS OF STEGANOGARPHY

RED PIXELS=397059 BLACK PIXELS=172129 WHITE PIXELS=2 BLUE PIXELS=0 GREEN PIXELS=0

(c) Enter Password (d) Description of Pixels

ENTER PASSWORD TO REVEAL MESSAGE :-

234567

MESSAGE

YOUR MESSAGE IS ----russians

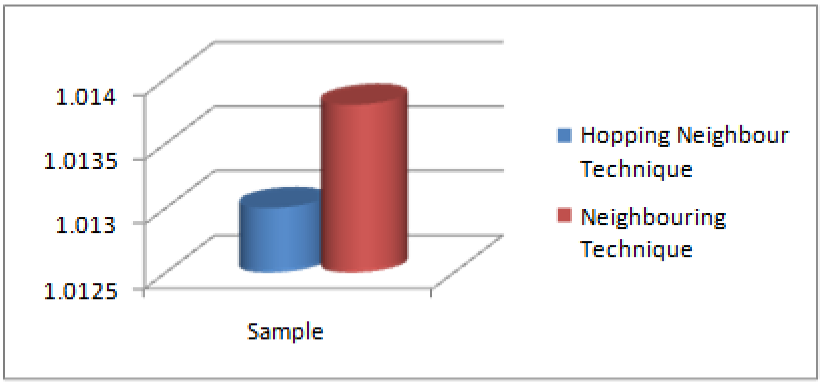
(e) Reveal Password (f) Reveal Information

Figure.1: Sample Image

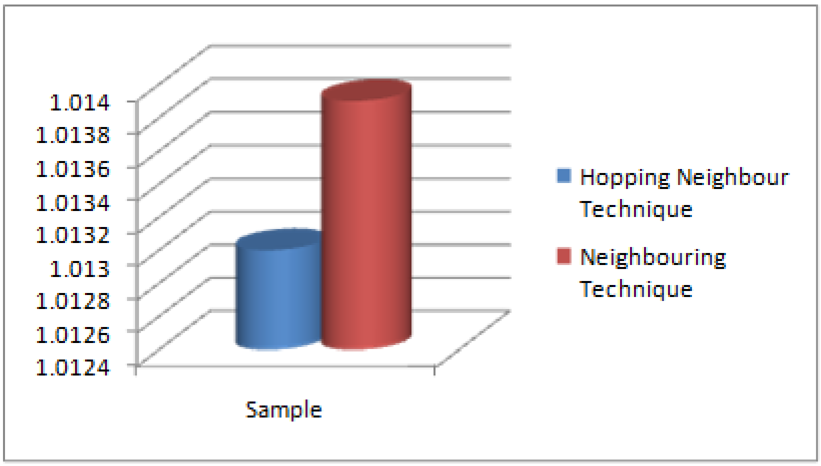
Comparison Table

Parameters	Hopping Neighbour Technique	Neighbouring Technique
Hiding Capacity	397059	99264.75
Time Taken	1.0130	1.0139
Linear Visual Distortion	0.2604	1.0417

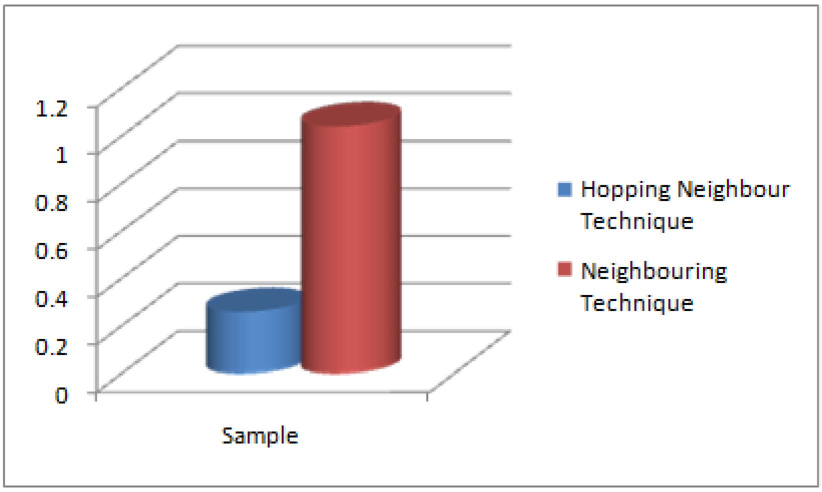
Comparison Graphs



Graph.1: Comparison of Data Hiding Capacity



Graph.2: Comparison of Time taken to hide information



Graph.3: Comparison of Linear Visual Distortion

V. CONCLUSION & FUTURE SCOPE

Conclusion

The research work is concluded as the main reason of using this technique is Simplicity, it is easily understandable to the user and there are negligible changes in the image after embedding the secret information. Significant amount of data can be hidden i.e. hopping neighbour technique provides larger capacity of hiding the information than the neighbouring approach. Largest cluster is chosen for the purpose of hiding the information so that enough information can be embedded without degrading the quality of image. So that it does not arouse the suspicion of the intruder if the quality degrades. Hopping Neighbour technique is one of the most efficient technique to hide the information in the image. This technique provides minimal linear visual distortion in image as in comparison with neighbouring technique after embedding the information into the image. Visual distortion is less because the the largest cluster is chosen and information is hidden in least influential cluster of the image.

Future Scope

The work that is performed in this thesis can be used for future research in various areas which are as following:
This project work can be extended in future for the research purpose by applying this method on 3D images. The hopping neighbour technique is implemented on .jpeg image file format and can be used over .bmp, .tif, .png file formats in future. The process can be fastened up if we are able to compress the image without loss of information. The future research can be carried on the robustness of JPEG images, which are able to resist attacks like cropping, blurring or resizing of the image. The process of hiding the data requires the human intervention and it future research can be done to automate the initialization process. This work can also be tested over larger data set of images. Now I am working on hiding data in Images with help of hopping neighbour technique but in future my work can be extended to other media like video or audio files. This technique can also be compared with other techniques and along with different parameters to prove its efficiency.

ACKNOWLEDGMENT

I would like to express my gratitude to my guide Mr Vikas Wasson for his continuous support and guidance. He kindly read my paper and offered invaluable detailed advices on grammar, organization and the theme of the paper.

REFERENCES

[1] Gonazalez R., Woods R. et al. , " Digital Image processing" 3rd edition ,2007.

[2] Yang C.,Liu F. "Steganalysis Frameworks of Embedding in Multiple Least-Significant Bits" *IEEE Transactions On Information Forensics And Security,* Vol. 3,pp. 4, 2008.

[3] Roy R.,Changder S. "Evaluating Image Steganography Techniques: Future Research Challenges" *IEEE International Conference On Computing Management And Telecommunication*, pp. 309-314, 2013.

[4] Bohme.R, "Principles of Modern Steganography and Steganalysis" Advanced Statistical Steganalysis, Springer Berlin Heidelberg, pp. 11-77, 2010.

[5] Cheddad A.,Condell J. "Digital image Steganography: Survey and analysis of current methods" *Elsevier Signal Processing* Vol. 90, pp. 727–752,2010.

[6] Judge J. "Steganography: Past, Present, Future", SANS Institute Publications, 2001.

[7] Mathkour H.,Assassa G. "A Novel Approach for Hiding Messages in Images" *IEEE International Conference on Signal Acquisition and Processing,* pp. 89-93,2009.

[8] Chandramouli R.,Memon S.*"*Analysis Of LSB Based Image Steganography Techniques", *IEEE International Conference on Image processing*, Vol. 3, pp. 1019-1022, 2001.

[9] Khamrui A.,Mandal J."A Genetic Algorithm based Steganography using Discrete Cosine Transformation (GASDCT)" *Elsevier Procedia Technology* vol.10 pp.105 – 111, 2013.

[10] Majumder A.,Changder S. "A Novel Approach for Text Steganography: Generating Text Summary using Reflection Symmetry" *Elsevier International Conference on Computational Intelligence: Modeling Techniques and Applications,*Vol. 10, pp. 112-120,2013.

[11] Begum M.,Venkataramani Y. "LSB Based Audio Steganography Based On Text Compression" *Elsevier International Conference on Communication Technology and System Design,*Vol. 30, pp. 703-710,2011.

[12] Marvel L.,Boncelet C. "Spread Spectrum Image Steganography" *IEEE Transactions On Image Processing,* Vol. 8,pp. 1075-1083,1999.

[13] Katzenbeisser S.,Petitcolas F., " Information Hiding Techniques for Steganography and Digital Watermarking*",* *Artech House,* Boston, Vol. 2, 2004.

[14] Song S.,Zhang J. "A Novel Secure Communication Protocol Combining Steganography and Cryptography" *Elsevier Procedia Engineering* Vol. 15,pp. 2767 – 2772, 2011.

[15] Altaay A.,Sahib S."An Introduction To Image Steganography Techniques"*IEEE International Conference On Advanced Computer Science Aplications And Technologies*, pp. 122-126, 2012.

[16] Maya S.,Miyatake M. "Robust Steganography using Bit Plane Complexity Segmentation" *IEEE 16th International Conference on Electrical and Electronics Engineering*, pp.51-51, 2006.

[17] Raman G.,Subhalakshmi R. "Active Steganalysis based on Adapted Lempel-Ziv complexity and Approximate Entropy Estimation", Proceedings of 2013 *IEEE Conference on Information and Communication Technologies,* pp. 917-918, 2013.

[18] Mathkour H.,Sadoon B. "A New Image Steganography Technique" *IEEE International Conference On Wireless Communications, Networking And Mobile Computing*, pp.1-4, 2008.

[19] Sullivan K.,Madhow U. "Steganalysis for Markov Cover Data With Applications to Images" *IEEE Transactions On Information Forensics And Security,* Vol. 1,pp. 2,2006.

[20] Ashwin S.,Kumar S. "Novel and Secure Encoding and Hiding Techniques using Image Steganography: A Survey" *IEEE International Emerging Trends In Electrical Engineering And Energy Management*, pp.171-177, 2012.

[21] Artz D. 'Digital Steganography: Hiding data within Data', *IEEE Internet Computing,* vol. 5, pp.75-80, 2001.

[22] AkhtarN.,Khan S."An Improved Inverted LSB Image Steganography" *IEEE International Conference on Issues and Challenges in Intelligent Computing Technologies*, pp. 749-755 , 2014.

[23] Biswas D.,Biswas S." Digital Image Steganography using Dithering Technique" *Elsevier Procedia Technology, Elsevier* Vol. 4 , pp. *251 – 255,* 2012.

[24] Cao H.,Kot A. “On Establishing Edge Adaptive Grid for Bilevel Image Data Hiding”, *IEEE Transactions On Information Forensics And Security*, Vol. 8,2013.

[25] Chanu Y.,Tuithung T.“A Short Survey on Image Steganography and Steganalysis Techniques”, *IEEE National Conference On Emerging Trends And Application In computer Science*, pp. 52-55, 2012.

[26] Garg M., Wasson V. “Data Security With Image Clustering Using Hopping Neighbour Technique” *International Journal of Emerging Research in Management and Technology,* Vol. 3, pp. 142-145,2014.

[27] Luo X.,Wang D. “A review on blind detection for image Steganography” *Elsevier Signal Processing,*Vol. 88,pp. 2138-2157, 2008.

[28] Noda H.,Paulding J. “Application of Bit-Plane Decomposition Steganography to JPEG2000 Encoded Images” *IEEE Signal Processing Letters*, Vol. 9,pp. 12,2002.

[29] Provos N.,Honeyman P. "Hide and Seek: An Introduction to Steganography." *IEEE Security & Privacy* , Vol. 3, pp. 32-44, 2003.

[30] Roy S., Venkateswaran D. “A Text based Steganography Technique with Indian Root” *Elsevier Procedia Technology* Vol. 10, pp.167 – 171, 2013.