



An Improved Approach of Text Steganography in Application with Rotational Symmetry

B.Ramapriya

Assistant Professor, Dept. of Computer Science, Sri Akilandeswari women's College, Wandiwash, Tamil Nadu, India

ABSTRACT: Steganography is the art and science of covered writing that involves in communicating secret message in an appropriate multimedia channel. The purpose of Steganography is secret communication to hide the existence of a message from the snooping eyes. The goal is to conceal the existence of the embedded data. Steganography has various useful applications. It has been propelled to the vanguard of current security techniques by the remarkable growth in the computational power. Steganography ultimate aim is undetectable, robustness and capacity to hide data and separate it from the relevant techniques with watermarking and cryptography Digital algorithms have been developed by using texts, images and audio as the wrap media. However, using text as the target medium is relatively difficult as compared to the other media, because of the lack of redundant information in a text file. This paper presents an approach for text Steganography through a technique that uses rotational symmetry of the English alphabets. To hide secret data bits, the proposed method checks the rotational symmetry of both vertical and horizontal properties of the characters present in sentence of the text. If it is followed, it selects the sentence to generate a summary of the text, known as cover text or stegotext. Correspondingly at the extraction, the receiver checks for the rotational symmetry properties followed by the characters present in the sentences of the stegotext and places the matching bits to get the secret message from the summary generated by the hiding process. The proposed method shows a satisfactory result with the cover text chosen from different newspapers.

KEYWORDS: Digital Text Steganography, Rotational Symmetry, Summary, Security

I.INTRODUCTION

Steganography is the art of hidden writing its presence cannot be detected [1]. The aim of Steganography is to hide secret message from the third party. For decades people strive to develop innovative methods for secret communication. A thorough history of Steganography was found in the literature [2-4]. Three techniques are interlinked, Steganography, watermarking and cryptography. The first two are difficult to dominate especially for those coming from different area. Steganography differs from cryptography, the art of hidden writing, which is intended to make a message unreadable by an involuntary receiver but does not conceal the existence of the secret communication. Even though Steganography and Cryptography are different and distinct, these two can be treated as twin sisters of secret communications. As the application of computer in real life is increasing every day, the need to secure data is becoming more and more essential and challenging part of data transfer and therefore the hidden exchange of information have attracted more attention to the researchers. The secret message is encoded in such a way that information's survival kept hidden from the unintended receivers. The aim of Steganography is to establish a secured communication in an absolutely unnoticeable manner [5] and to avoid drawing distrust to the transmission of a hidden data [6].

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirce.com

Vol. 5, Issue 7, July 2017

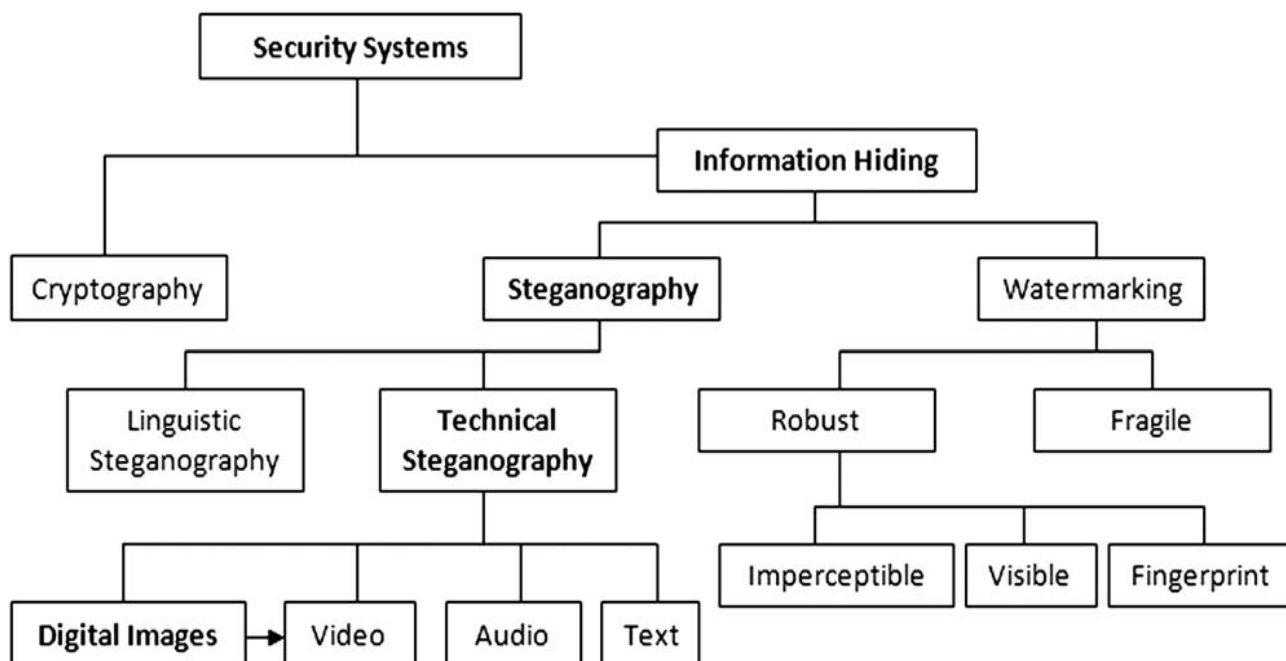


Fig.1 The different disciplines of information hiding in steganography

The word steganography is originally derived from Greek words which mean “Covered Writing”. It has been used in various forms for thousands of years. In the 5th century BC Histaiacus shaved a slave’s head, tattooed a message on his skull and the slave was dispatched with the message after his hair grew again [2–4]. It was also reported that the Nazis invented various Steganography methods during World War II such as Microdots, and have reused invisible ink and null ciphers. In 1945, Morse code was concealed in a drawing (see Fig. 2). The secret information is encoded onto the stretch of grass beside the river. The long grass denoted a line and the short grass denoted a point. The decoded message read: “Compliments of CPSA MA to our chief Col Harold R. Shaw on his visit to San Antonio May 11th 1945” [7].

Steganography is engaged in various useful applications such as copyright control of materials, enhancing robustness of image search engines and smart Identity cards where individuals details are entrenched in their photographs, video–audio synchronization, companies safe circulation of secret data, TV broadcasting, TCP/IP packets[2], checksum embedding [8], Petitcolas [9] established some contemporary applications, one of which was in Medical Imaging Systems where a separation is considered necessary for confidentiality between patients’ image data or DNA sequences and their captions, e.g., physician, patient’s name, address. This should be maintained between the two. Thus, embedding the patient’s information in the image could be a useful safety measure and helps in solving such problems. Steganography would provide a crucial guarantee of approval that no other security tool may guarantee.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirce.com

Vol. 5, Issue 7, July 2017

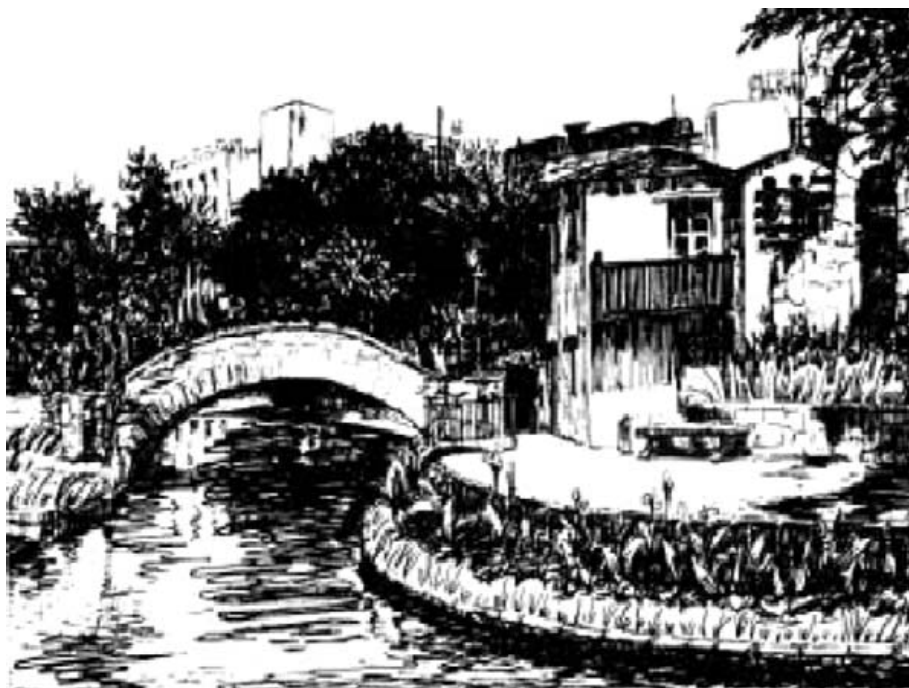


Fig.2. Suppression of Morse Code with hidden message (1945)

The Steganography is not to keep others from knowing the secret information, but it is to keep others from drawing misgiving that the information even exists. If a Steganography method causes someone to expect that there is secret information in a carrier medium [10]. The first written support about Steganography being used to send messages is the Heredotous [11] story about slaves and their shaved heads. The modern Steganography can be given in terms of the prisoner's problem [12]. Let Alice wishing to send a secret message S to Bob. In order to do so, Alice embeds S into a cover-object C to obtain the stego-object C . The stego-object C is then sent through the public medium. In a Steganography framework, the technique for embedding the message is unknown to Wendy and shared as a secret between Alice and Bob. However, it is generally not considered as good practice to rely on the secrecy of the algorithm itself. In private key Steganography Alice and Bob share a secret key which is used to embed the message.

Many Steganography methods have been introduced on different cover media such as images [5, 10, 13], video files [14,15] and audio files [16]. Due lack of large scale repeated information in a text file when compared with images, audio and video files. Text Steganography seems to be most complicated kind of Steganography [17]. This paper presents a novel approach for text Steganography by generating the summary of a text file that contains English language text. The proposed method takes as input a widely available text and the secret message. The secret message is hidden in the summary by following the rotational symmetry properties of the characters of English alphabets along the axis of rotation. As result of the system, a review is generated from the chosen input text and that review is our cover text. That review is to be sent to the receiving end. At the other end depending on the same properties of the alphabets, relevant secret bits from the cover text are generated to get back the Input message.

II. RELATED WORKS

A. Lexical Steganography

In lexical Steganography lexical units of natural language text such as words are used to cover secret bits. In this approach a word could be replaced by its synonym and the choice of word to be chosen from the list of synonyms would be based upon secret bits. These transformation should preserve the grammatics of the sentence. In this method we refer to a legitimate object in a communication as a cover object, while a message with embedded information is



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirccce.com

Vol. 5, Issue 7, July 2017

called a stego object. The stegosystem creates stego-objects with embedded messages, using a stego-key, such that the operation cannot be inverted insignificantly. For example consider a sentence – Teresa is a good lady. If good represent 00 then according to the input bits 01, 10, 11 we can replace the word good by nice, pleasant and kind correspondingly to hide the bits.

B. *Ontological Steganography*

In Ontological Steganography, to embed message instead of perfectly leaving semantics unbroken by replacing only synonymous words an explicit model for the meaning is used to evaluate equivalence between texts. This method is also having the disadvantage like NICETEXT that sometimes it may create semantically erroneous texts.

C. *Syntactical Steganography*

Bearing in mind, the syntactic structures of a text, the syntactical Steganography approach construct syntactically correct sentences by using the Context Free Grammars (CFG). CFG based Mimicry [12], NICETEXT [13] comes under this category. NICETEXT uses the cover text as a source of syntactic patterns and by running the cover text through a part-of-speech tagger. This algorithm obtains a set of "sentence frames," e.g. [(noun) (verb) (prep) (det)(noun)] for 'He saw in the Room'. Then by using the dictionary, a large list of (type, word) pairs where the type may be based on the part-of-speech tagger or its synonyms. The system randomly generates series of words to form a sentence. Though, NICETEXT produces syntactically correct sentences, the output text is almost always set of ungrammatical and semantically inconsistent sentences. Using this disadvantage of NICETEXT steganalysis algorithms [14] have been developed to find the presence of secret information in the cover file generated by NICETEXT.

D. *Audio Steganography*

Audio Steganography is a technique used to transfer hidden information by modifying an audio signal in an invisible manner. It is the science of hiding some secret text or audio information in a host message. The host message before Steganography and stego message after Steganography have the same personality. Sampling is the process in which the analogue values are only captured at regular time intervals. Quantization converts each input value into one of a discrete value.

E. *Image Steganography*

Image Steganography is a technique used to transfer hidden information by modifying an image signal. It has many lossy compression techniques to embed data in the quantized DCT coefficients of JPEG images. In PNG image Steganography it modify the LSB each pixel in the PNG.

F. *Other Techniques*

In Text Steganography by Hiding Information in particular Character of Words [18] approach, specific characters from particular words are selected to hide the message. For example, the first character of every alternative word hides the secret information. Text Steganography by Line Shifting method [19, 20] is useful approach where lines are shifted vertically to some degree. For example, lines are shifted vertically to degree say α or $-\alpha$. For α , the information is 1 and for the 0 information is $-\alpha$. This method is suitable for printed text. Information can be hidden by Creating Spam Texts [21] in a HTML file.

This approach uses the suppleness of HTML regarding case-sensitiveness. By Word Shifting method [19,22], information is hidden in the text by shifting words horizontally and by altering the distance between the words. Feature Coding method [23, 24] changes the feature or structure of the text to hide data. For example, elonging or shortening end portion of some characters, or by vertical displacement of points of characters like 'i', 'j' etc. In this method a large volume of data can be hidden in the text.

By adding Open Spaces method [25], the information can be hidden by adding extra white spaces in the text. In addition all these, some algorithms for text Steganography through Indian Languages have been proposed by using feature coding method [26] and dynamic programming method [27]. Some more approaches in text Steganography area has also been developed, that scans the letters in English alphabets and review their shapes for hiding secret data [28, 29]. Some Techniques are also there that checks the properties of sentences and depending on how it hides secret data in it [30].



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

III. PROPOSED WORK

In most of the text Steganography algorithms the secret message is concealed by changing the structure of the cover file, hence there is a chance of either distrust or loss of data in case of retyping. To claim more security, in this proposed method, instead of hiding the secret bits by changing the structure of the text file, we hide the secret message by generating summary of a given text collected from public media. The process of generating the summary is dependent on the rotational symmetry property of English alphabets and according to that, they are grouped into different sets, where each set represents a pair of bits. To do this let us first analyze the rotational symmetry properties followed by the alphabets and classify them to represent the binary bits.

A. Classification of Alphabets

For classification of the English Alphabets following rotational symmetry property, first we select the horizontal axis as the axis of symmetry and divide the English letters into two groups based on the horizontal bisection of the letters i.e. whether a character is equal or not on both side of the axis after bisection of it e.g. characters like 'A', 'C' are not same on both side of axis if bisected horizontally, whereas the letters 'B', 'H' etc. are just the reverse i.e. same on both side of axis after horizontal bisection. The vertical bisection of symmetry classify the English letters into two groups.

Group ID	Group name	Letters in Group	Bits to be hidden
1	Reflection Property followed along Neither axis	C, F, G, J, L, N, P, Q, R, Z	00
2	Reflection Property followed along Horizontal Axis	B, D, E, K, S	01
3	Reflection Property followed along Vertical Axis	A, M, T, U, V, W, Y	10
4	Reflection Property followed along Both Axis	H, I, O, X	11

Table.1 Groups based on bisection of English letters

B. Hiding Algorithm

Steps:

1. Convert the surreptitious data to binary bit stream.
2. Check, if the whole length of the bit stream is even or odd. If odd, add an extra '0' bit in the MSB. Now separate the total bit stream in the groups of two bits each.
3. Convert the text file to upper case.
4. For every group of secret bits, check whether any of the character that correspond to the group following Table 1, is the starting character of the sentence, provided the first word is not an article. In second case, go to the first letter of the next word.
 - (a) If agreed, select the sentence and add it to the cover text.
 - (b) If not, move to the next sentence.
5. The process is to be carried out, until the total secret bit stream is exhausted.
6. The result cover text is the generated summary of the text and is to be sent to the intended recipient.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

Output:

Cover Text, The obtained binary bit stream is converted back to the alphanumeric form, to get the secret message

C. Extraction Algorithm

Steps:

1. Scan the first letter of the first word of each sentence, provided the first word is not an article. In the second case, scan the first letter of the subsequent word.
2. Verify, the obtained character belongs to which group according to Table 1. Append the bit value, represented by the character to a file.
3. Convert the surreptitious bit stream to the alphanumeric form.
4. The obtained alphanumeric note is our established secret message. Output:
 1. Secret Message

IV. EXPERIMENTAL RESULTS

The text file we have used for execution purpose is shown in Fig. 3. For implementing the algorithms we have selected the secret message to be hidden in the cover text, as "Hi", as shown in Fig. 4. To hide the message the program first converts the secret message to its corresponding binary stream '00010111001011'. Since the length of the bit stream is even, so we divide the total bit stream into groups of two bits each and the obtained first group is '00'. We have used the approach based on bisection of the letters along both horizontal and vertical axis for implementation and for that purpose we have used the Table 1. Accordingly the system starts scanning the selected text, shown in Fig. 3 and the first sentence obtained is "Going green seems to be the new experiment in some college laboratories who have decided to take a novel, eco-friendly route to conserve energy." The first letter of the first word of the scanned sentence is 'G', that belongs to the group '00' as after bisection of 'G' along both vertical and horizontal axis, we find that in both the case the two bisected halves are not equal. As 'G' is capable of hiding '00', so the corresponding sentence would be selected and saved in the cover file i.e. the file containing summary of the text.

Following the same method scanning from left to right, the next sentence the letter starts with 'R' which comes under the bits 00, so the bit stream given is 01. It should not be included in the cover file. We must move on to the next sentence. At the receiver end the input is the cover file.

The first scanned sentence is "Going green seems to be the new experiment in some college laboratories who have decided to take a novel, eco-friendly route to conserve energy". The first letter is G the corresponding bits are 00. It should be included in the file. This way, we obtain the bit stream '00010111001011'. After that, the generated bit stream is converted to alphanumeric values considering each 7 bits at a time. The characters are generated that was the original secret message. Hi that was hidden in the text of the Fig.6

Going green seems to be the new experiment in some college laboratories who have decided to take a novel, eco-friendly route to conserve energy. Replacing Bunsen burners with cigarette lighters and toxic chemicals with safer alternatives are a few of the green measures .Paul Wilson, head of chemistry at Madras Christian College, says going green is important. Industries that have a report on green measures under category 7 of the National Assessment and Accreditation Council, academic institutions are also rated on the eco-friendly practices they adopt," he says. He points out that precautionary measures need to be inculcated in faculty and Students. For example, a Bunsen burner need not be used for all chemical reactions; instead using cigarette lighters can help save LPG. He adds, More than 50%energy can be saved as some experiments require minima; heating.

Fig.3 Text from newspaper



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

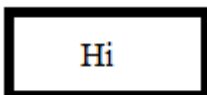


Fig.4 Secret Text to be hidden

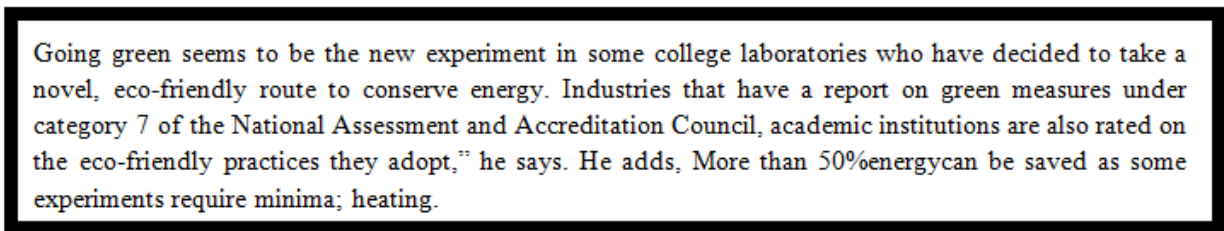


Fig.5 Cover Text

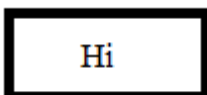


Fig.6 Extracted Secret message

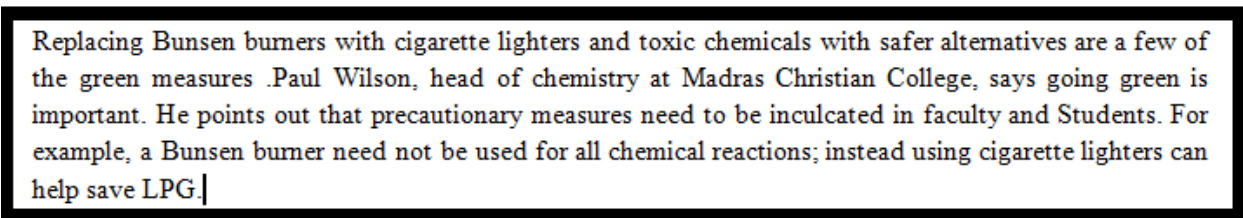


Fig.7 Summary of the selected Text

In the same way, Cheddad et al.[31] proposed a security technique which protects scanned documents from forgery using this techniques. This method not only points out forgery but also allows forensic experts to gain access to the original document despite being manipulated in Fig.7

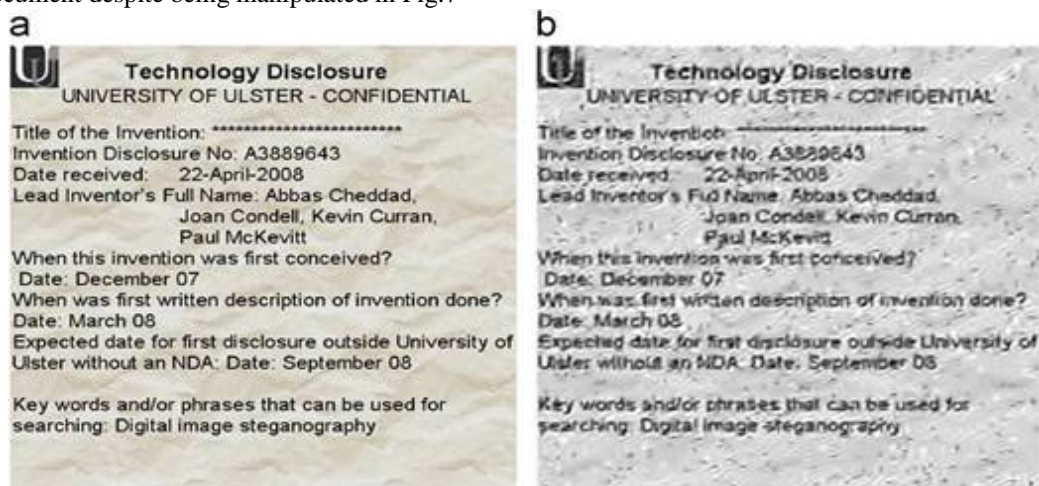


Fig.7. digital document forgery detection (a) Stego-image (b)

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 7, July 2017

The Fig.7. (a) contains the document scanned to find the forgery. Fig.7. (b) contains the stego-image contains the secret message, The Fig.7. (c) contains the attacked stego image of the document. The Fig.7.(d) contains the inverse of half toning of the hidden image. The Fig.7.(e)contains the error signal. The Fig.7. (f) contains the document after thresholding operation



Fig.7. (c) attacked stego-image, (d) inverse half toning of hidden data

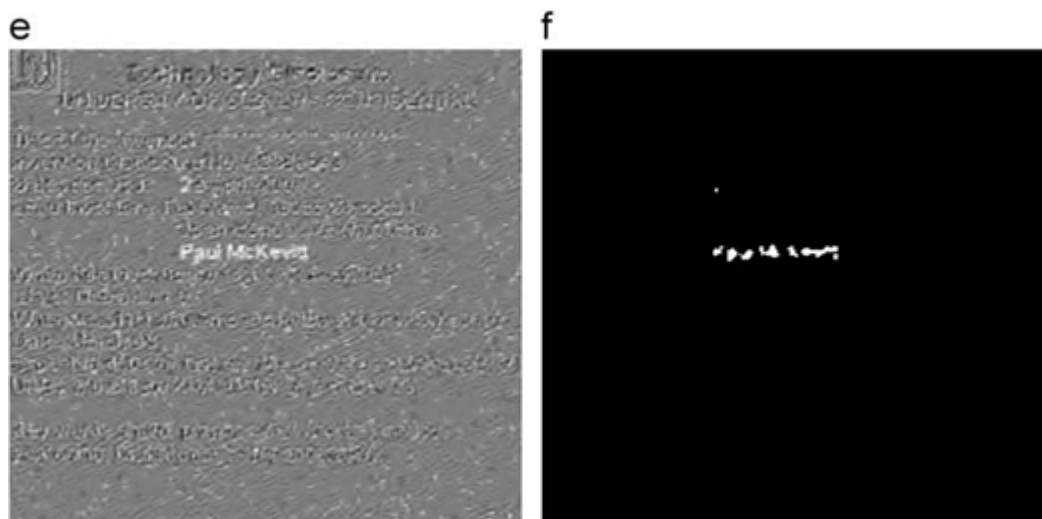


Fig.7. (e) error signal, (f) after thresholding operation

V. CONCLUSION AND FUTURE SCOPE

In this paper we have introduced an improved approach for text Steganography by generating cover file and secret message using the rotational symmetry of the alphabets. To hide secret data, the proposed method checks the rotational symmetry like horizontal, vertical and inverse rotational properties present in the Text. The generated summary is the text that is the cover file, this stego-text is generated by the system. Similarly, at the receiver end, the receiver checks



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirccce.com

Vol. 5, Issue 7, July 2017

for the rotational symmetry followed by the characters present in the sentence and places the subsequent bits to get the secret message. This paper also discusses the types of Steganography. Further research can be made to increase the security from the prying eyes. And also increase the robustness of the system

REFERENCES

- [1] C . Cachin, "An Information-Theoretic Model for Steganography", in proceeding 2nd Information Hiding Workshop, vol. 1525, pp. 306-318, 1998
- [2] N.F.Johnson,S.Jajodia,Exploring steganography: seeing the unseen, IEEE Computer 3 vol.12,pp.26–34, 1998.
- [3]J.C.Judge, Steganography: past, present, future. SANS Institute publication,/http://www.sans.org/reading_room/whitepapers/steganography/552.phpS, 2001.
- [4] N.Provos,P.Honeyman,Hide and seek:an introduction to steganography, IEEE Security and Privacy 1,Vol.3,pp. 32–44, 2003.
- [5] R Chandramouli, N. Memon, "Analysis of LSB Based Image Steganography Techniques", IEEE pp. 1019-1022,2001.
- [6] N.F. Johnson, S. Jajodia, "Staganalysis: The Investigation of Hiding Information", IEEE, pp. 113-116, 1998.
- [7] J.P.Delahaye,Informationnoy_ee, informationcach,PourlaScience /www.apprendre-en-ligne.net/crypto/stega no/229_142_146.pdfS (in French),Vol.229 pp.142–146, 1996..
- [8] W. Bender, W. Butera, D. Gruhl, R. Hwang, F.J. Paiz, S. Pogreb, Applications for data hiding, IBM Systems Journal 39 Vol.3, pp. 547–568, 2000.
- [9] F.A.P. Petitcolas, Introduction to information hiding, in: S. Katzen- beisser, F.A.P. Petitcolas (Eds.), Information Hiding Techniques for Steganography and Digital Watermarking, Artech House, Inc., Norwood, 2000.
- [10] D. Artz, "Digital Steganography: Hiding Data within Data", IEEE Internet Computing, pp. 75-80, May-Jun 2001.
- [11] Herodotus. The Histories. Penguin Books, London. Translated by Aubrey de Séincourt, 1996.
- [12] G. Simmons, "The prisoners problem and the subliminal channel," CRYPTO, pp.51–67, 1983.
- [13] J. Chen, T. S. Chen, M. W. Cheng, "A New Data Hiding Scheme in Binary Image," in Proc. Fifth Int. Symp. on Multimedia Software Engineering. Proceedings, pp. 88-93, 2003.
- [14] G. Doërr and J.L. Dugelay, "A Guide Tour of Video Watermarking", Signal Processing: Image Communication, vol. 18, Issue 4, pp.263-282, 2003.
- [15] G. Doërr and J.L. Dugelay, "Security Pitfalls of Frameby- Frame Approches to Video Watermarking", IEEE Transactions on Signal Processing, Supplement on Secure Media, vol. 52, Issue 10, pp. 2955-2964, 2004.
- [16] K. Gopalan, "Audio steganography using bit modification", Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing,(ICASSP '03), vol. 2, 6-10, pp. 421-424, April 2003.
- [17] J.T. Brassil, S. Low, N.F. Maxemchuk, and L.O'Gorman, "Electronic Marking and Identification Techniques to Discourage Document Copying", IEEE Journal on Selected Areas in Communications, vol. 13, Issue. 8, pp. 1495-1504, October 1995.
- [18] P. Wayner, "Mimic functions", Cryptologia XVI, pp. 193–214, July 1992.
- [19] M. T. Chapman, "Hiding the hidden: A software system for concealing ciphertext as innocuous text", Master's thesis, University of Wisconsin-Milwaukee, May 1997.
- [20] Peng Meng, Liusheng Huang, Zhili Chen, Wei Yang, Dong Li," Linguistic Steganography Detection Based on Perplexity", International Conference on MultiMedia and Information Technology, pp 217-220, 2008.
- [21] T.Moerland,"Steganography and Steganalysis", May 15, 2003, www.liacs.nl/home/tmoerlan/privtech.pdf.
- [22] Y. Kim, K. Moon, and I. Oh, "A Text Watermarking Algorithm based on Word Classification and Inter-word Space Statistics", Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR'03), pp. 775–779, 2003.
- [23] K. Rabah, "Steganography-The Art of Hiding Data", Information Technology Journal, vol. 3, Issue 3, pp. 245-269, 2004.
- [24] Shirali-Shahreza, M.H.; Shirali-Shahreza, M., "A New Approach to Persian/Arabic Text Steganography ".Computer and Information Science, 2006. ICIS-COM SAR 2006. 5th IEEE/ACIS International Conference on 10-12 July, Page(s):310 – 315, 2006.
- [25] D. Huang, and H. Yan, "Interword Distance Changes Represented by Sine Waves for Watermarking Text Images", IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 12, pp. 1237-1245, December 2001.
- [26] S. Changder, N.C. Debnath , "An Approach to Bengali Text Steganography", Proceedings of the International Conference on Software Engineering and Data Engineering (SEDE-08), ISBN: 978-1-880843-67-3, pp. 74-78, Los Angeles, California, USA, July, 2008,.
- [27] S. Changder, N.C. Debnath, D. Ghosh, "LCS based Text Steganography through Indian Languages" Proceedings of 2010 3rd IEEE International Conference on Computer Science and Information Technology(ICCSIT 2010), ISBN: 978-1-4244-5539-3,pp. 53-58(vol 8) July, 2010.
- [28] Shradha Dulera, Devesh Jinwala, Aroop Dasgupta, "Experimenting with the Novel Approaches in Text Steganography" International Journal of Network Security & Its Applications (IJNSA), Vol.3, No.6, November 2011
- [29] S. Changder, S.Das, D.Ghosh, "Text Steganography through Indian Languages using Feature Coding Method", Proceedings of the 2nd International Conference on Computer Technology and Development (ICCTD), 2010, 10.1109/ICCTD.2010.5645849, Page(s): 501 – 505, November 2010.
- [30] S. Changder, N.C. Debnath, D. Ghosh, "A Greedy approach to Text Steganography using Properties of Sentences" Proceedings of the Eighth International Conference on Information Technology: New Generations (ITNG),2011, ISBN: 978-1-61284-427-5 Pages(s): 30-35, April, 2011.
- [31] A. Cheddad, J. Condell, K. Curran, P. Mc Kevitt, A secure and improved self-embedding algorithm to combat digital document forgery, Signal Processing Vol.89,pp. 2324–2332, 2009.