

Available online at www.sciencedirect.com**ScienceDirect**

Procedia Technology 10 (2013) 167 – 171

Procedia
TechnologyInternational Conference on Computational Intelligence: Modeling Techniques and Applications
(CIMTA) 2013

A Text based Steganography Technique with Indian Root

Souvik Roy^{a,*}, P.Venkateswaran^b^aJadavpur University, India^bJadavpur University, India

Abstract

Steganography is the art and science of covered writing. Modern digital steganography uses text, images, audio, video etc. as a cover medium. This paper presents a text based steganography technique based on the Vedic Numeric Code. Frequency of the letters in English alphabet in conjunction with Vedic Numeric Code are used for the steganography technique. No separate importance is given for vowels and consonants.

© 2013 The Authors. Published by Elsevier Ltd. Open access under [CC BY-NC-ND license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Selection and peer-review under responsibility of the University of Kalyani, Department of Computer Science & Engineering

Keywords: Steganography; Text Steganography; Information Security; Vedic Numeric Code

1. Introduction

Steganography is the art of hiding of a message within another so that presence of hidden message is indistinguishable. The key concept behind steganography is that message to be transmitted is not detectable to the casual eye. This is also the advantage of steganography over cryptography. An unhidden encrypted message, no matter how unbreakable, raises suspicion but steganography is used not to keep other from knowing the hidden information, but it is to keep other from knowing the hidden information [1]. There many steganography methods which use images [2, 3], video [4], audio [5] as a cover media.

Text steganography uses text as a cover media for hiding message. Message can be hidden by shifting word and line [6, 7], in the open spaces [8], in word sequence [9]. Properties of a sentence such as number of words, number of characters, number of vowels, position of a vowel in a word are also used to hide secret message. The advantage of preferring text steganography over other steganographic technique is its smaller memory requirement and simpler

* Corresponding author. Mob: +919433420386

E-mail address: souvice31@gmail.com.

communication [10]. But due to lack of large scale redundancy of information in text file, in compared to other medias, text steganography seems to be most difficult kind of steganography [11].

This paper presents a different approach to the English text based steganography with Indian root. In the propose method, no properties of a sentence are used rather characteristics of the English language is used. This gives flexibility and freedom from the point view of the sentence construction but it increases computational complexity.

2. Related work

In [1] by using the properties of a sentence and the presence of redundant feature code able characters in Indian Languages, a message is hidden into an innocent cover file containing Indian texts. Using the existence of too many points in Persian and Arabic phases, information is hidden in the Persian and Arabic texts [12]. Hindi letters and its diacritics and numerical code are used in [13] for hiding message into Hindi text. Generating a random sequence of characters or words, specific information can be hidden in sequence [9] but it often results in meaningless words or sentence which is susceptible to raise suspicion. In method [14], some specific characters from certain words, placed in certain sequence in sentence, are used as hiding place for secret message but the method is time consuming and takes a lot of mental power. By placing punctuation signs in specific places, information is hidden [9] but information hiding capacity of the method is low. Using synonym of certain words in a sentence, information can be hidden [15]. By altering the features of a text information is hidden in text [16].

3. Suggested Algorithm

Some of the outstanding characteristics, inflexion, fixed word order and use of periphrases, of the English language are used for the steganography technique. Inflexion means that it can indicate the relationship of the words into a sentence with a minimum change of shape. In fixed order, the place of each word in a sentence decides its relationship with the others. Periphrases are the different ways to express something.

Sri Bharati Krishna Tirthaji in [17] described a particular code called Vedic Numerical Code used in deciphering Sanskrit text. The coding is based on the tongue position. For applying the Vedic code to the English alphabet, frequency of letters in English vocabulary [18] is used as the basis of assigning numbers to the letters in the English alphabet. Frequency of letters is shown in table 1. No discrimination is made for assigning coding number to vowels and consonants as compared to [13].

Each letter in the alphabet is assigned a number in the range of 0 to 15 as shown in table 1.

3.1. Encoding

For the proposed text based steganography technique, each letter in the secret message is represented by its ASCII code and obtained ASCII code is expressed in 8 bit binary number. The 8 bit binary number is then divided into two 4 bit parts. Each 4 bit part, representing a number in the range 0 to 15, is then used to choose corresponding suitable letters from table 1. A meaningful sentence is constructed by using letters obtained as the first letters of suitable words. To give flexibility in sentence construction, articles, pronoun, preposition, adverb, was/were, is/am/are, has/have/had, will/shall, would/should are not taken into consideration for encoding during sentence construction.

3.2. Decoding

For decoding the cover message, first letter of each word is taken and represented by the corresponding 4 bit binary number and these 4 bit binary numbers are then combined to obtain 8 bit binary numbers and corresponding ASCII codes for letters in the secret message are then obtained from the 8 bit numbers. Finally the secret message is recovered.

3.3. Table

Table 1. Table of letter frequency and number assignment

Letter	Frequency of letter	Number assigned	Letter	Frequency of letter	Number assigned
E	11.1607 %	15	M	3.0129 %	7
A	8.4966 %	14	H	3.0034 %	7
R	7.5809 %	13	G	2.4705 %	6
I	7.5448 %	13	B	2.0720 %	5
O	7.1635 %	12	F	1.8121 %	4
T	6.9509 %	11	Y	1.7779 %	4
N	6.6544 %	11	W	1.2899 %	3
S	5.7351 %	10	K	1.1016 %	3
L	5.4893 %	10	V	1.0074 %	3
C	4.5388 %	9	X	0.2902 %	2
U	3.6308 %	8	Z	0.2722 %	2
D	3.3844 %	8	J	0.1965 %	1
P	3.1671 %	7	Q	0.1962 %	0

4. Experimental Result

To implement the above text based steganography method, a secret message is considered. Suppose it is “text”. The message has 4 letters and by representing the letters in ASCII code and converting it into binary form, 32 bits are obtained. The result of encoding is shown below.

Promod Yadav has gone to Bangalore for the marriage of his daughter to Pormash Yadav.

Here first letter of each word is used for encoding. Encoding is not case sensitive. In another example, secret message is “Steganography”. Encoded result is shown below.

Byomkesh Kumar Hazra, father of Govindo Byomkesh of Gobordanga Pally in Golpukur Jela, has given an application to George’s executive group for a medical history of Zygomycosis in Govindo, Joint Medical Group also urgently placed it to the committee.

5. Information Security

A brief description about the application of above technique in information security during online shopping is given. The steganography technique in conjunction with visual cryptography technique [19] can provide a secure online shopping experience by ensuring customer data privacy in terms of providing only limited information both to the shopping service provider and payment gateway but still ensuring smooth fund transfer. An authorization code such as password can be hidden inside a cover text and a snapshot of the cover text is taken. Now from the snapshot two shares are generated by using (2,2) visual cryptography scheme. One of the shares is kept by the bank and other share is given to the customer. Now during fund transfer in online shopping, submission of its share by the

customer to the bank side ensures fund transfer from the customer account to the shopping service provider account without providing any credit or debit card information to the third party. Visual cryptography ensure the security of the authentication data during from customer side to certified authority (CA) side from where only the needed information is given to shopping service provider side and security of the rest of the information from the CA side to the bank side is provided by steganography technique in conjunction with traditional cryptography technique like Triple DES, RC4 etc. Fig.1 is the snapshot of the cover message, Fig.2 is generated share 1, Fig.3 is generated share 2 and Fig.4 is the reconstructed cover message.

Promod Yadav has gone to Bangalore for the marriage of his daughter to Pormash Yadav.

Fig.1.

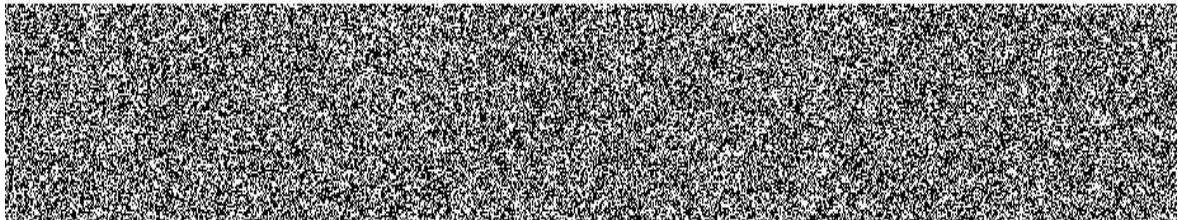


Fig.2.

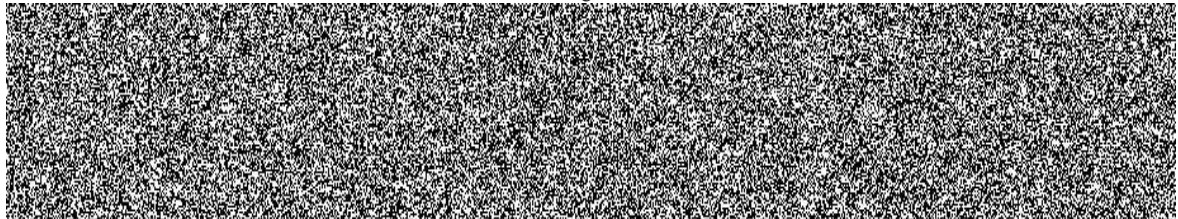


Fig.3.

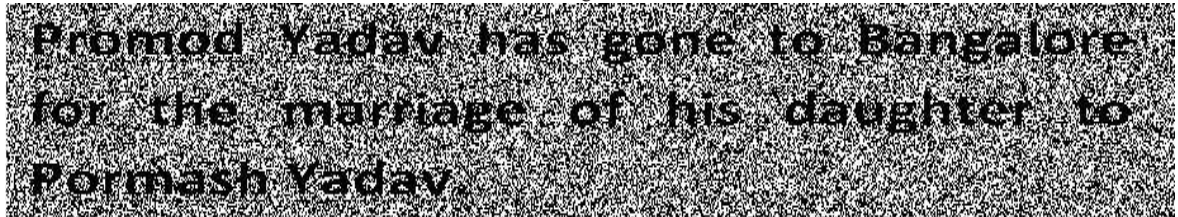


Fig.4.

6. Conclusion

As compared to [13], which gives separate importance to vowels and consonants, this technique does not give any separate importance to vowels and consonants providing better flexibility in hiding data in case of English language. As certain words are not included during encoding to create better cover medium, results in large number of words as compared to [13]. In the first example, to hide a 4 letter word, 8 words are required excluding the words that are added to provide flexibility in sentence construction. In the second example, 25 words are required for hiding a 13 letter word. So to hide a larger message, this technique requires large no of words and creates a complexity in sentence construction. However in banking and online shopping application where one required to send only account no and password, this disadvantage of the technique is used in its advantage by creating spam messages to hide one's account no and password or any other personal information and avoid rising suspicion. In case of mobile banking, transaction can be made via SMS securely using the proposed technique.

In conjunction with other cryptography technique like visual cryptography, the proposed text steganography technique can provide a two layer of authentication and security system in physical and online banking as well as online shopping as shown in section information security.

References

- [1] S.Changder, N. C. Debnath, D.Ghosh, “A Greedy Approach to Text Steganography using Properties of Sentences”, Proceedings of the 2011 Eighth International Conference on Information Technology(ITNG 2011), ISBN: 978-0-7695-4367-3, pp. 30-35, Las Vegas, NV, USA.
- [2] R Chandramouli, N. Memon, “Analysis of LSB Based ImageSteganography Techniques”, Proceeding of the 2001 International Conference On Image Processing, vol.3, pp. 1019-1022, 2001.
- [3] J. Chen, T. S. Chen, M. W. Cheng, “A New Data Hiding Scheme in Binary Image”, Proceeding of Fifth International Symposium on Multimedia Software Engineering, pp. 88-93, 2003.
- [4] Hu ShengDun, U. KinTak, “A Novel Video Steganography Based on Non-uniform Rectangular Partition”, Proceeding of 14th International Conference on Computational Science and Engineering, pp. 57-61, Dalian, Liaoning, 2011.
- [5] Daniel Gruhl, Anthony Lu, Walter Bender, “Echo Hiding”, Proceedings of the First International Workshop on Information Hiding, pp. 293-315, Cambridge, UK, 1996.
- [6] Jack Brassil, Steven Low, Nicholas Maxemchuk, Larry O’Gorman, “Hiding Information in Document Images”, Proceedings of the 1995 Conference on Information Sciences and Systems, Johns Hopkins University, 1995.
- [7] Jack Brassil, Steven Low, Nicholas Maxemchuk, Larry O’Gorman, “Electronic Marking and Identification Techniques to Discourage Document Copying”, Proceedings of IEEE INFOCOM’94,vol.3, pp. 1278-1287, Toronto, June 1994.
- [8] Walter Bender, Daniel Gruhl, Norishige Morimoto, A. Lu, “Techniques for Data Hiding”, IBM Systems Journal, Vol. 35, 1996.
- [9] K. Bennet, “Linguistic Steganography: Surevey, Analysis, and Robustness Concerns for Hiding information in Text”, Purdue University, Cerias Tech Report 2004—2013.
- [10] J.C. Judge, “Steganography: Past, Present, Future”, SANS Institute, November 30, 2001.
- [11] 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.
- [12] M. Hassan and Shirali Shahreza “A New Appraoach to Persian and Arabic Text Steganography”, Fifth IEEE/ACIS International Conference on Computer and Information Science (ICIS COMSAk’06), pp. 310-315., July 2006.
- [13] Kalavathi Alla, Dr. R. Siva Rama Prasad, “An Evolution of Hindi Text Steganography”, Proceeding of Sixth International Conference on Information Technology, pp. 1577-1578, Las Vegas, NV, 2009.
- [14] T. Moerland, “Steganography and Steganalysis”, Leiden Institute of Advanced Computing Science.
- [15] M. Niimi, S. Minewaki, H. Noda, and E. Kawaguchi, "A Framework of Text-based Steganography Using SD-Form Semantics Model", Pacific Rim Workshop on Digital Steganography 2003, Kyushu Institute of Technology, Kitakyushu, Japan, July 3-4, 2003.
- [16] K. Rabah, "Steganography-The Art of Hiding Data", Information Technology Journal, vol. 3, Issue 3, pp. 245-269, 2004.
- [17] Bharati Krishna Tirthaji, “Vedic Mathematics and its Spiritual Dimension”, Motilal Bansari Publishers,1992.
- [18] <http://oxforddictionaries.com/words/what-is-the-frequency-of-the-letters-of-the-alphabet-in-english>.
- [19] Moni Naor, Adi Shamir, “Visual Cryptography”, EUROCRPT 1994.