Impact Factor:					
	ISRA (India) = 4.971	SIS (USA) = 0.912	ICV (Poland) = 6.630
	ISI (Dubai, UAE) = 0.829	РИНЦ (Russia) = 0.126	PIF (India) = 1.940
	GIF (Australia) = 0.564	ESJI (KZ) = 8.997	IBI (India) = 4.260
	JIF = 1.500	SJIF (Morocco) = 5.667	OAJI (USA) = 0.350

SOI: 1.1/TAS DOI: 10.15863/TAS

International Scientific Journal

Theoretical & Applied Science

p-ISSN: 2308-4944 (print) **e-ISSN:** 2409-0085 (online)

Year: 2020 **Issue:** 07 **Volume:** 87

Published: 22.07.2020 **http://T-Science.org**

QR – Issue QR – Article

A.A. Ganiev
Tashkent University of Information Technologies
researcher, Tashkent, Uzbekistan

O.N. Mavlonov
Samarkand branch of Tashkent University of Information Technologies
Researcher, Samarkand, Uzbekistan
mavlonov_obit@mail.ru

THE ANALYSIS OF TEXT STEGANOGRAPHY METHODS

Abstract: *Today, in any area of our life, any type of data (text, image, audio, video, etc.) can be stored and transmitted at high speed. However, they are easily accessed illegally, forged, and copied. For this reason, the problem of information security has become more important with the development of the computer. One of the grounds discussed in the field of information security is the exchange of information through mass media using steganography methods. In this article, we discuss various approaches to text steganography. There are several methods of text steganography, each type has its special function, and they all have their strengths and weaknesses. We analyze some of the main approaches of text steganography and compare them depending on their effectiveness based on advantages and disadvantages.*

Key words: *steganography, ciphertext, encrypt, hash, cover object, attackers, cryptography, stegotext.*

Language: *English*

Citation: Ganiev, A. A., & Mavlonov, O. N. (2020). The analysis of text steganography methods. *ISJ Theoretical & Applied Science, 07 (87),* 85-88.

Soi: http://s-o-i.org/1.1/TAS-07-87-23 ***Doi***: https://dx.doi.org/10.15863/TAS.2020.07.87.23

Scopus ASCC: *1700.*

Introduction

Steganography is a branch of information concealment, and its main purpose is to securely transmit or transmit data in a completely undetectable way. The literal meaning of writing in the cover is the practice of hiding messages in other messages to hide the existence of the original[1,3]. The inventor of the word " steganography "is Trithemius, the author of early publications on cryptography: "Polygraphy and steganography". The technical term itself comes from the Greek word steganos, meaning "covered", and graphia, meaning "written". Steganography is the art of hidden communication [4]. The very existence of the message is secret. In addition to invisible ink, an often-cited example of steganography is the ancient story of Herodotus, who tells of a slave sent by his master Hist To the Ionian city of Miletus with a secret message tattooed on his head. After the tattoo, the slave grew his hair to hide the message. Then he went to Miletus and on arrival shaved his head to show the message to the Regent of the city, Aristagoras. The message prompted Aristagoras to start a revolt against the Persian king. In this scenario, the message is of primary importance to the Hist, and the slave is simply the message carrier [4].

We can classify the steganography methods based on the covers in the following manner: text, image, audio, video, Protocol.

Text steganography

Text steganography is a method of using written natural language to hide a secret message. In-text documents, we can hide information by making changes to the structure of the document without making noticeable changes to the corresponding output. Storing a text file requires less memory, and its faster and easier communication makes it preferable over other types of steganographic methods.[2] Text steganography can be broadly classified into three types:

- Format based
- Random and Statistical generation

- Linguistic methods

Format based methods

This method uses physical text formatting as a space to hide information. Inserting spaces or unreadable characters, serious Tin errors throughout the text and changing the font size are some of the many formatting methods used in text steganography. Some of these methods, such as deliberate spelling mistakes and inserting spaces, can deceive readers who ignore random spelling mistakes, but can often be easily detected by a computer [5].

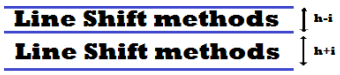

Figure 1. Line Shift method

Random and statistical methods

To overcome the problem of comparing with known plain text, steganographers often prefer to create their accompanying texts. The methods used are hiding information in a random sequence of characters, and statistical properties of word length and letter frequency are used to create words that will have the same statistical properties as actual words in a given language.

Linguistic methods

Linguistic steganography, in particular, considers the linguistic properties of the generated and modified text and in many cases uses the linguistic structure as a space in which messages are hidden [7, 8].

Table 1. Advantages and disadvantages of Text steganography methods

Text Steganography methods	Advantage	Disadvantage
Line Shift	This method is suitable only for printed text.	When attackers use OCR (character recognition program), the hidden information easily would get destroyed.
Word Shift	Word shifting method identifies less because of the change of the distance between words to fill line is quite common.	If someone knows the algorithm of distances, using the difference in the distances one can obtain the hidden text by comparing the stego text with the algorithm. Also, retyping or using OCR programs destroys the hidden information.
White Steg	Since in practically all text editors, extra white space at the end of lines is skipped over, it won't be noticed by the casual viewer.	Inconsistent use of white space is not transparent.
Semantic method	Attackers cannot detect by retyping or using OCR programs.	The smart reader who has a huge knowledge of words can discover their synonyms or antonyms.
Syntactic method	The amount of information hidden behind the method is trivial.	It requires the identification of correct places to insert punctuation marks.
CSS	Using RSA public-key cryptosystem and ciphertext makes it more secure.	Text Correlation Program or any function corrected text is easily detected.
Mixed-case font	The hiding capacity will be very high compared to other text steganography methods.	Attackers can easily detect the special program. Retyping and using OCR programs destroys the hiding information.

SMS-Texting	Attackers cannot detect by retyping or using OCR programs.	It takes a long time to make a wordlist and the capacity of the text must be large to hide the secret message.
Feature Coding	A large volume of information can be hidden in the text without making the reader aware of the existence of such information in the text.	By placing characters in a fixed shape, the information is lost. Retyping the text or using the OCR program destroys the hidden information.
SSCE (Secret Steganographic Code for Embedding)	Using the SSCE table and a certain mapping technique by inserting articles (a, an) with the nouns increases the security of the data.	Inserting articles would attract smart reader, especially, when non-specific nouns are used a lot.
Hiding data in the wordlist	It is based on the special calculated algorithm. The secret message is hidden to the text without any changes and cover is dynamically generated.	If attackers would be aware of the algorithm of this method, it is easy to detect it.
Hiding data in paragraphs	The approach works by hiding a message using the start and end letter of the words of a cover file. A word having the same start and end letter is skipped. Since no change is made to cover, the cover file and its corresponding stego file are the same.	The volume of data hiding in the paragraph would be very less. The capacity of hiding the large volume of data leads to the challenge.

CONCLUSION

There are several methods of text steganography that are analyzed and provided with the advantages and disadvantages of each method (Table 1). Each method has its special algorithm, the corresponding ability to hide data in the text, and the use of a sphere, which makes it more secure. Using the line offset method, we can hide a huge amount of data, but the line offset method is only intended for printed text, because in this method, except for the printed text reorganization program (OCR), the hidden information is destroyed. In the syntactic method (.) And (,) are used to send very important information and hide a very small amount of data. However, a smart reader can easily detect or destroy a secret message.[9] The semantic method is effective and its security is higher than in the previous method. In my opinion, hiding data in paragraphs is the most effective way to securely transfer confidential data over the Internet.

References:

1. Zaynalov, N. R., Narzullaev, U. K., Muhamadiev, A. N., Bekmurodov, U. B., & Mavlonov, O. N. (2019). “Features of using Invisible Signs in the Word Environment for Hiding Data,” *Int. J. Innov. Technol. Explore. Eng.*, vol. 8, no. 9S3, pp. 1377–1379, DOI: 10.35940/ijitee.i3295.0789s319.
2. Rasulovich, Z. N., Nuralievich, M. A., B. U. B. ugli, Nizomovich, M. O., Utkirovich, K. J., & Dusmurod, Q. (2019). *“Information Security Issues For Travel Companies,”* in 2019 International Conference on Information Science and Communications Technologies (ICISCT), 2019, pp.1–4, DOI: 10.1109/ICISCT47635.2019.9011896.
3. Islomov, S. Z., Mavlonov, O. N., Muhamadiev, A. N., Shodmonov, D. A., & Djumaev, S. N. (2018). “New authentication scheme for cloud computing,” *J. Adv. Res. Dyn. Control Syst.*, vol. 10, no. 10, pp. 2316–2319.
4. Cox, I.J., Miller, M.L., Bloom, J.A., Fridrich, J., & Kalker, T. (2000). *“Digital watermarking and*

steganography". Morgan Kaufman Publishers, Second edition.

5. Cummins J., Diskin, P., Lau, S., & Parlett, R. (2004). *"Steganography and Digital watermarking"*. School of computer science, The University of Birmingham.
6. Kumar, K.A., Pabboju, S., & Shyam, N.M. (2014). "Advance text steganography algorithms: an overview", *IJRA*, Vol. 1, No. 1(1), pp. 31-35.
7. Por, L.Y., & Delina, B. (n.d.). *"Information Hiding: A New Approach in Text Steganography".*
8. Saraswathi, V., & Kingslin, S. (2014). "Different approach to text steganography: a comparison", *IJERMT*, Vol. 3, No. 11, pp. 124-127.
9. Por, L.Y., Ang, T.F., & Delina, B. (2008). "WhiteSteg - a new scheme in information hiding using text steganography", *WSEAS Transactions on Computers*, Vol.7, No.6, pp. 735-745.
10. Bhattacharyya, S., Banerjee, I., & Sanyal, G. (2010). "A Novel Approach of Secure Text-Based Steganography Model using Word Mapping Method(WMM)", *International Journal of Computer and Information Engineering*, Vol.4, No. 2.
11. Por, L.Y., Ang, T.F., & Delina, B. (2008) WhiteSteg-a new scheme in information hiding using text steganography. *WSEAS Trans Comput* 7(6):735-745.
12. Shirali-Shahreza, M. H., & Shirali-Shahreza M. (2006). *A new approach to Persian/Arabic text steganography*. In: 5th IEEE/ACIS international conference on computer and information science and 1st IEEE/ACIS international workshop on component-based software engineering, software architecture, and reuse, (pp.310-315).
13. Krishnan, R. B., Thandra, P. K., & Sai Baba, M. (2017). *An overview of text steganography. 4th International Conference on Signal Processing, Communications and Networking* (ICSCN - 2017), March 16 - 18, 2017, Chennai, INDIA
14. Bender, W., Gruhl, D., Morimoto, N., & Lu, A. (1996). Techniques for data hiding. *IBM Syst J* 3(3&4):313-336.
15. Shirali-Shahreza, M. H., Shirali-Shahreza, M. (2008). *A new synonym text steganography*. International conference on intelligent information hiding and multimedia signal processing, (pp. 1524-1526).
16. Brassil, J.T., Low, S.H., Maxemchuk, N.F., & O'Gorman, L. (1995). *Document marking and identification using both line and word shifting*. Proceedings of INFOCOM '95 proceedings of the fourteenth annual joint conference of the IEEE computer and communication societies, (pp. 853-860).
17. Cummins, J., Diskin, P., Lau, S., & Parlett, R. (2004). S*teganography and digital watermarking.* School of Computer Science, (pp. 1-24).
18. Blinova, E.A. (2016). The steganographic method based on the line-shift coding method on non-displayed symbols of the electronic text document. Belarusian State Technological University. *BGTU*, № 6, pp.166-169.
19. Kabetta, H., & Dwiandiyanta, B.Y. (2011) Suyoto: information hiding in CSS: a secure scheme text steganography using public-key cryptosystem. *Int J Cryptogr Inf Secur* 1(1):13-22.
20. Simmons, G.J. (1984). *The prisoner`s problem and the subliminal channel*, Proc. Workshop on Communications Security (Crypto`83), 51-67.