An Acquaintance to Text-Steganography and its Methods

A P Singh¹, S Moudgil² and S Rani³
¹Computer Science Engg, SGT University, Gurugram, India
²Computer Science Engg, SGT University, Gurugram, India
³Computer Science Engg, SGT University, Gurugram, India
E-mail: aadips11996@gmail.com

Abstract. In this research paper discussion will be going on upon text steganography. As the word ‘steganography’ is very well known amongst human beings, considering that, this research paper will highlight the term ‘Text Steganography’. Defining steganography, that an individual can hide data into/within or behind text, images, audio, and video would not be inappropriate. Digital steganography is a modern approach and necessity of digital security system. Digital Text steganography is a distinct factor of digital steganography. Various types of Text steganography will have been discussed further. Keywords- Steganography, Text Steganography, Information Hiding, Linguistic.

1. Introduction
The word stego means covered/hidden writing and the origin is ‘Greek Language’. Steganography itself is an art and a science of communication by hiding the presence of information underneath some fixed media resources like Text, Audio, Images and Video [1]. Here the journey begins to understand the ‘Security System’ over ‘Computer Networking’. In the security system there are two parts of it Cryptography and Steganography. Cryptography works upon alphabetical letters means upon words. Whenever an individual is intended that he/she wants to send a message through a network but not in a readable manner there cryptography takes place to change the format of message being sent on a network. So that no third person other than intended ones could read or understand the message. Since this research paper is all about ‘Text Steganography’ keep cryptography aside. Now start the discussion about steganography, information is considered as an essential and important asset of humankind [2]. Steganography helps to hide the information within a cover media being transferred on a network. Table no. 1 shows difference between Cryptography and Steganography and Figure 1 gives the general idea about steganography mechanism.

| Cryptography                           | Steganography                  |
|---------------------------------------|--------------------------------|
| Message is visible but not in readable/understandable manner | Information is hidden in cover media |
| Most popular and widely used          | Less popular                   |
| It means secret writing               | It means covered writing       |
| study of attacks in cryptography is called | Study of attacks in Steganography is called |
| Cryptanalysis                         | Steganalysis                   |
| Less use of mathematical concept      | Use of Mathematic in almost every technique |
1.1 Steganography:
Information Hiding, overcome trespassing, hiding information from Hackers are some main goals of steganography. Some essential and fundamental objectives of steganography are below:

- Information to be hidden: That secret message that you don’t want anyone to know about and ready to be embedded, is first bet to make a stego file;
- Chosen media to cover information: The media object one of audio, image, video or text is called as cover media, beneath the confidential information is hidden;
- Obtained mixture-object: The final and resultant state of whole process is obtained mixture-object. Often this state is called as stego file/stego object;
1.2 Text Steganography:
One of the components of steganography is text steganography, and this steganography is most arduous kind because in it there is no unnecessary information as compared with image and a sound file [10]. Text steganography has been divided into three types:

- Format based methods
- Linguistic methods
- Random and statistical generation methods

Format based methods: Deliberate misspelling, resized fonts, spaces insertion, hidden characters are some various styles adopted in format-based methods. In this method text has been used as a physical location in which the information is hidden. Space insertion, deliberate misspelling are some of those methods that can fool a humankind easily because as a human, human beings don’t point out the little mistakes often, but computer will detect those things immediately.

Linguistic methods: Linguistic properties of a text come under the linguistic steganography methods. The generated text has been used as space to hide the information.

An Example:

Random and statistical generation methods: Randomly concealing order of characters and words is considered in random and statistical generation methods. These methods are used for “automatically cover-text generation” by using the statistical properties of a language [11]. These are completely “order of characters and word” based methods [1] [2]. PCFG (Probabilistic context-free grammar) is a language model to be used to generate the stego text.

![Figure 3. Mechanism of Text Steganography](image-url)
2. Related Work

The methods of Text Steganography, every method has got its own pros and cons. Study of Text Steganography aims to overcome the difficulties of covering of important and vulnerable information as Akbar et al., have shown [1]. With the help of some tools and advancement of technology many kind of researches have been performed by researchers in the field of Text Steganography in the manner to enhance the Text Steganography. Sharma et al., researched that, withstand algorithms for attacks are very important and be a first priority in Text security [2]. To improve the efficiency of algorithms of steganography, there are several researchers working upon this research domain out there. Concealing the message in such fashion, attendance of hidden information could not be thought. Shivani et al., [3]. Wang et al., has proposed a methodology in their research paper named as “Emoticon-based Text Steganography in chat”. Proposed methodology can be combined with other methodologies as a hybrid methodology [4]. Evolution of mobile phones, computers, world wide networking with low cost and rich features has led mankind to a major security concern. SMS-Texting language in SMS steganography now has another steganography method by Rafat [5].

The illustration that steganography creates gives a good hope because of Its broadly transmission over world wide networking. Roy et al., [6]. Gupta et al., proposed a method “discrete wavelet transform based steganography” for concealing information into cover images, not having major changes [8].

3. Text Steganography Methods

- Line shift method
- Word shift method
- Syntactic method
- Abbreviation/Acronym method
- Semantic method
- Hiding data using white spaces method
- Feature coding method

**Line Shift Method:** In Line shifting method the lines of text are being shifted vertically thus information encoding has been performed. Shifted lines represents the bits of secret message.

| Shifts lines up slightly up or down | h | i |
|-------------------------------------|--|--|
| Lines to be shifted decided by codebook | h+i |

**Word Shift Method:** The words of a text into a line are shifted horizontally little bit to show the bits of altered information/secret message, still keeping the natural appearance of spacing.

**Syntactic Method:** In this method punctuations are used to hide the information within. Punctuation symbols like Comma (,) in a text and full stop (.) in a text are altered with secret information. If the reader of the text, that has secret information in it, has good knowledge about grammar of that particular language the secret information is covered with, can easily discover the secret message easily. This method of text steganography is easy in comparison of other steganography methods.
Abbreviation/Acronym method: The short form of any word is used instead the whole word to hide the 0s and 1s of secret message.

**Table 2. Abbreviation/Acronym method**

| Abbreviation/Acronym | Actual word          |
|----------------------|----------------------|
| ASAP                 | As Soon As Possible  |
| L8                   | Late                 |
| KH                   | Kill Him             |
| V                    | We                   |
| B2B                  | Body to Body         |

Semantic Method: In this method synonyms of words are used to hide the binary bit pattern of information.

**Table 3. Semantic method**

| Answer | Reply       |
|--------|-------------|
| Begin  | Initiate    |
| Calm   | peaceful    |
| Cry    | Weep        |
| End    | Finish      |
| Funny  | Humorous    |

**Table 4. Change of spelling in American and British language**

| American Standard of English Language | British Standards of English language |
|--------------------------------------|---------------------------------------|
| Flavor                              | Flavour                               |
| Humor                               | Humour                                |
| Labor                               | Labour                                |
| Color                               | Colour                                |
| Fiscal                              | Financial                             |
| Center                              | Centre                                |

Hiding data using white space method: Secret information is represented by space between words, paragraphs and sentences. Below, an appropriate example of white space method is shown.

```
T h e  q u i c k  b r o w n  f o x
```

NORMAL TEXT
Feature coding method: colour of the text, Font of the text are some features of the text. Features of a text is being used in this method to hide the bits of secret information. What feature is going to be altered it depends upon code word.

E.g. This line is an example of feature coding method

4. Conclusion
This theoretical paper gives a Brief information about Text-Steganography and shows how each method of Text-Steganography is distinct from each other. An individual can easily understand the methods, discussed above in this paper because of the simplicity of the language and given examples along with each method. The Text-Steganography plays an important role amongst Information-Hiding system even amongst other kinds of steganography such as Audio-Steganography, Image-Steganography and Video Steganography. It has been proved that Text could be used as a cover media to conceal the confidential information. World has gotten almost 6,500 languages, and if computers were working in 6,500 languages then, the Cryptography and Steganography would have been more difficult and complex.

References
[1] Akbar F C et al, A Study of Text Steganography Methods, Journal of Engineering and Applied Science 15 (2): 369-372, ISSN: 1816949X, © Medwell Journals, 2020.
[2] Sharma S et al, Analysis of Different Text Steganography Techniques: A Survey, 2016 Second International Conference on Computational Intelligence & Communication Technology, 978-1-5090-0210-8/16 $31.00 © 2016 IEEE, DOI 10.1109/CICT.2016.34.
[3] Yadav S V K, Bathum S, A Novel Approach of Bulk Data Hiding using Text Steganography, 3rd International Conference on Recent Trends in Computing 2015 (IRCTC2015), Procedia Computer Science 57 (2015) 1401 – 1410.
[4] Wang Z H et al, Emoticon-based Text Steganography in Chat, 2009 Second Asia-Pasific Conference on Computational Intelligence and Industrial Applications, 978-1-4244-4607-0/09/$25.00 ©2009 IEEE, PACIIA 2009.
[5] Rafat K F, Enhanced Text Steganography in SMS, International Islamic University, Islamabad, Pakistan.
[6] Roy S et al, A novel approach to format based text steganography, Conference Paper – January 2011, DOI: 10.1145/1947940.1948046 Source: DBLP.
[7] Shirali-Shahreza M H et al, Text Steganography in Chat, 1-42441007-X/07/$25.00 ©2007 IEEE
[8] Gupta S et al, An Innovative Method of Text Steganography, 2015 Third International Conferencen Image Information Processing, 978-5-5090-0114-8/15/$31.00 © 2015 IEEE.
[9] Koley S, Mandal K K, A Novel Approach of Secret Message Passing Through Text Steganography, International Conference on Signal Processing, Communication, power and Embedded System (SCOPUS)-2016, 978-1-5090-4620-1/16/$31.00 © 2016 IEEE.
[10] Weldu T A, A Comparison of Different Approaches in Text Steganography, IJSR, ISSN (Online): 2319-7064, PaperID: ART2018487, DOI: 10.21275/ART2018487.
[11] Shirali-Shahreza M et al, Text Steganography in SMS, 2007 International Conference on Convergence Information Technology, 0-7695-3038-9/07 $25.00 © 2007 IEEE.