November-December 2019
ISSN: 0193-4120 Page No. 6030 - 6037

Hiding Message in Text Steganography using RGB Color in Random Location

[1]Baharudin Osman, [2]Azman Yasin, [3]Mohd Nizam Omar
[1,2,3]School of Computing, College of Arts and Sciences, University Utara Malaysia 06010,
[1]bahaosman@uum.edu.my

Article Info
Volume 81
Page Number: 6030 - 6037
Publication Issue:
November-December 2019

Article History
Article Received: *5 March 2019*
Revised: *18 May 2019*
Accepted: *24 September 2019*
Publication: *27 December 2019*

Abstract
Hiding data in text has always been a challenge due to textual media contains limited redundant space compares to the others media such as image, audio and video. This paper proposed a hiding technique of hidden message at random location using Pseudorandom Number Generator and RGB color. A hidden message has been represented in 3D representation by enhance the equation introduce by the previous study. The result shows that the capacity of hidden message achieves 100%. However the representation of hidden message character shows a similar pattern which can cause the steganalysis to study the pattern of concealment.

Keywords: *Text steganography, RGB color, Pseudorandom number generator, 3D representation*

1. Introduction

Steganography is one of the branches in the security system and it differs from cryptography where cryptography converts hidden messages into unreadable forms. However, this unreadable message is easier to attract compare to the steganography technique which hiding an information in a documents arouse any suspicious. Steganography is the art of concealing information with the aim of concealing data from third parties so that no one suspects the existence of a hidden message during transmission. The word steganography comes from the Greek (*steganos* + *graphy*) where *steganos* means "protected or hidden" and *graphy* means "writing". The first use of steganography was first introduced in 440 AD in the history of Herotodus where the message carrier would shave his hair and then the hidden message would be written on the scalp like a tattoo (Fridrich, 2010; Ingemar, Matthew, Jeffrey, Fridrich, & Kalker, 2008).

Steganography is more difficult and complex to attack by third parties compare to cryptography (Zielińska, Mazurczyk, &Szczypiorski, 2014) and the suspicion inherent in the message being sent can be eliminated (Bhat, Prabhu, & Renuka, 2017). Steganography is used in many fields such as military, medical diagnosis, business, financial and so on (Malik, Sikka, & Verma, 2017). Although steganography has been around since ancient times, according to Singh and Singh (2013), the number of researchers 'interested in steganography has increased and attracted researchers' attention and has evolved for two reasons. First, the

growth of the publishing and broadcasting industry has attracted particular parties to use data hiding techniques such as concealing copyright and serial numbers in digital films, audio recordings, books and multimedia products aimed at identifying the purity of a product. Second, there are also government and private sectors that do not allow encryption services to be used within the organization and make it a policy within the organization.

According to Sumathi, Santanam, & Umamaheswari(2013) steganography is a hot topic in the domain of information hiding and most studies focus on images, audio and video (Osman, Yasin, & Omar, 2016) as a medium of protection and less attention is given to the text medium. However, text is the most widely used medium today for daily data transmission (Bhaya, Rahma, & Al-nasrawi, 2013)and still has room to conduct research(Iyer, 2017). According to Ahvanooey, Li, Shim and Huang(2018) text is one of the major sources of data available and the most widely used digital media on the Internet, an important part of websites, books, articles, daily papers and so on. In addition, most important documents are in the form of texts such as appointment letters, certificates, reports, confidential documents and so on(Din & Utama, 2018).

2. Classification of Text Steganography

Text steganography can be broadly classified into three types as shown in namely, Format Based, Random and Statistical Methods, and Linguistic (Baawi, Mokhtar, & Sulaiman, 2018; Sharma, Gupta, Trivedi, & Yadav, 2016) as shown in Figure 1.

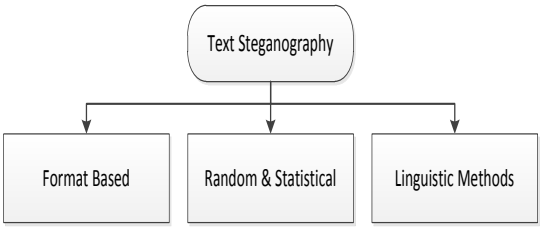

Figure 1: Steganography Methods

Format-based methods use a technique of manipulating blank or alphabetical space which are a popular methods used by most researchers. Linguistic method focuses on the manipulation of words that require a high level of knowledge of the language so that the generating text does not interfere with the semantics of the generated sentence. Finally, the statistical and random methods generate the stego text using techniques random and statistical techniques. Random techniques hide characters using random sequence replacement techniques while statistical techniques determine statistical values such as mean, variance, letter or word frequency, word length, etc. to determine the amount of information that can be hidden. This method will generate stego text based on the statistical properties of a document (Saraswathi & Kingskin, 2014). Baawi et al.(2018)classifies the cover text medium into five mediums as shown in Figure 2.2.

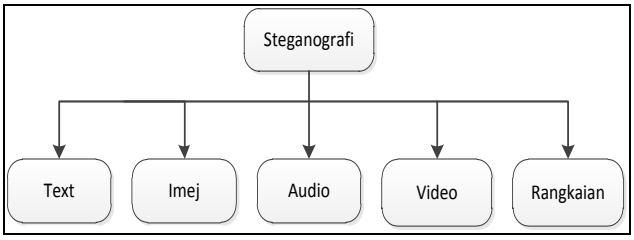

Figure 2: Cover Text Medium Classification(Baawi et al., 2018)

Figure 2.2 shows the cover medium used to hide hidden messages that can be classified into text, image, audio, video and network or protocol (Baawi et al., 2018; Krishnan,

Thandra, & Baba, 2017). Although these mediums are very popular (Hmood, Jalab, Kasirun, Zaidan, & Zaidan, 2010; Krishnan et al., 2017) and mostly used by researchers, but text medium is less focused by researchers due to the limited of redundant space to hide hidden messages(Bhattacharyya, Banerjee, & Sanyal, 2011; Krishnan et al., 2017; Sloan & Hernandez-castro, 2015).

However, text medium remains a researcher's choice as it uses less memory (Al-Azzawi, 2018) (Lwin & Phyo, 2014) (Shivani, Yadav, & Batham, 2015), fast delivery (Kumar & Sharma, 2014) and is less difficult to detect by steganalysis (Bhattacharyya, Indu, Dutta, Biswas, & Sanyal, 2011) compare to other mediums. This statement is supported by articles written by Zielińska et al.(2014) which stated that the two best features of the medium of protection are; first, the protective medium must be popular and; second, any changes to the physical object must not be detected or suspected by a third party

3. Related Works

A message can be hidden in cover text using properties attributes such as font, color and underline. Information hiding based on text attributes such as fonts, fonts and underlines are some of the techniques that researchers use in text steganography. According to Chaudhary and Dave(2016), conceal the hidden message based on text attributes can increase the capacity of hidden messages. Attributes such as the font type, font size, font style, font color can be used for the hiding process purpose. As such, RGB color attributes have been used in steganography studies by several researchers such as color change (Al-Asadi & Bhaya, 2016; Singh & Diwakar, 2014), character color and underline(Wang & Li, 2014), blank space color, margins and paragraph color (Stojanov, Mileva, & Stojanovi, 2014)and a combination of character and underline colors (Tang & Chen, 2013)using various RGB color values ranging between (0,0,0) to (255,255,255).

According to Khairullah(2019), the color technique used in email address for the concealment process in a study conducted by Malik et al.(2017)was identified having major weaknesses in terms of color used. In addition, the other issue identified is the sequencing of location sequences that are facilitated by steganalization techniques to manipulate the position of the hidden message characters. These factors have contributed to the researchers' research on this technique by focusing on hidden capacities and random hiding locations to enhance the performance of the generated text. According to Chaundhary, Dave and Sanghi(2016), concealing a hidden message based on a text attribute can increase the hiding capacity.

The ability to conceal hidden messages at random locations within the cover text can increase the level of steganography difficulty during extracting hidden messages. Random numbers play an important role in encryption especially in various network-based security applications (Elmahi, Wahbi, & Sayed, 2017). Random Number Generator (RNG) can be divided into three types: True Random Number Generator (TRNG), Pseudorandom Number Generators (PRNG) and Pseudorandom Number Function-(PNF). According to Srikanth, Mehta, Yadav, Singh, & Singhal (2017) PRNG approach is widely used for generating random numbers due to speed and has a recurrence relation where new values are generated based on previous values.

November-December 2019
ISSN: 0193-4120 Page No. 6030 - 6037

4. Proposed Method

This paper proposed a technique to embed a hidden message (HM) into cover text (CT) using a RGB color in a random location. Each HM character will be covert into ASCII code and then convert it to (x,y,z) or 3D representation to mapped with RGB color. Koley and Mandal(2017)technique convert an OCTAL value to (x,y) representation using equation (1).

$$(x * (x+1)/2) + y) \qquad (1)$$

The obtain value is then convert to ASCII code to map with a normal character. However the proposed method enhanced the technique by adding *z* axis into the 2D representation to form a 3D representation as shown in the following table.

Table 1: 3D Representation

Hidden Message (HM)	Koley & Mandal, (2017) (x,y)	Proposed (x,y),z
M	(14,10)	(14,10),0
E	(14,00)	(14,00),1
E	(14,00)	(14,00),2
T	(15,04)	(15,04),3
Y	(15,11)	(15,11),4
O	(14,12)	(14,12),5
U	(15,05)	(15,05),6
A	(13,10)	(13,10),7
T	(15,04)	(15,04),8
T	(15,04)	(15,04),9
E	(14,00)	(14,00),10
N	(14,11)	(14,11),11

The value of z will increase linearly when the size of hidden message increase with the maximum value is 255 due to the maximum RGB color code. The value represented by the proposed method will be map with the RGB color to format a character in cover text at a selected location using PRNG random number generator. PRNG use the following formulae (elmahni, 2017)

$x_{n+1} = (ax_n + t) \% m$; n= 0,1,2,3……

a,t,m: Integer constants

a. Embedding Process

Table 2 show the (*x,y,z*) value mapping with the random location of hidden message "MEETYOUATTEN" with the value of a,t,m,x_0 is 11,5,743,750 respectively.

Table 2: Mapping RGB with Random Location

Hidden Message (HM)	Proposed (x,y,z) (R,G,B)	PRNG Location
M	(14,10,0)	82
E	(14,00,1)	164
E	(14,00,2)	323
T	(15,04,3)	586
Y	(15,11,4)	507
O	(14,12,5)	381
U	(15,05,6)	481
A	(13,10,7)	95
T	(15,04,8)	307
T	(15,04,9)	410
E	(14,00,10)	57
N	(14,11,11)	632

The character at location 82, 164, 323, 586, 507, 381, 481, 95, 307, 410, 57, 632 will be format with the respective RGB color. For example, character 'r' and 'n' at location 82 and 164 will be formatted with RGB color (14, 10, 0) and (14,0,1) as shown in Figure 3.

November-December 2019
ISSN: 0193-4120 Page No. 6030 - 6037

In the research area of text steganography, algorithms based on font format have advantages of great capacity, good imperceptibility and wide application range. However, little work on steganalysis for such algorithms has been reported in the literature. based on the fact that the statistic features of font format will be changed after using font-format-based steganographic algorithms, we present a novel support vector machine-based steganalysis algorithm to detect whether hidden information exists or not. This algorithm can not only effectively detect the existence of hidden information, but also estimate the hidden information length according to variations of font attribute value.As shown by experimental results, the detection accuracy of our algorithm reaches as high as ninety nine point three percent when the hidden information length is at least sixteen bits.

Figure 3: Stego Text

Figure 3 shown, the generated stego text after conceal a hidden message.

b. Extraction Process

The extraction process will be execute by getting the value of *a*,*t*,*m* and x_0. The PRNG number is calculate to find a random location. The RGB color of that location will be extract to get the (x,y,z) representation and then convert to (x,y) representation using equation 1. The value of z shown the sequence character of hidden message. For example, with the value of a=11, t=5,m=743 and x_0 = 750, the random location value is 82, 164,323,586,507,381 …..The first character of location 82 will be extract the RGB color and the value of R and G is mapping with the equation 1 to get an original value which is

$$(14 * (14+1)/2) + 10) = 115$$

The value of 115 is an octal value which will convert to ASCII with the value is 77 and equivalent with "M" character.

5. Result

The performance of steganographic methods can be compared using three parameters such as capacity, imperceptibility and robustness. Capacity is the amount of hidden message can be concealing in a cover text and calculate by the following equation. The maximum capacity of hidden message is depend on the maximum size of hidden message can be conceal in cover text.

$$Capacity = \frac{amount\ of\ hidden\ message}{size\ of\ cover\ text}$$

Imperceptibility (invisibility) of the stego text is measure by determine the ability to detect the presence of hidden data and can be analyse by comparing the similarity of cover text and stego text using Jaro-Winkler Distance. The similarity score equal to 1 means both cover text and stego text are exactly same (Ahvanooey, Li, Hou, Rajput, & Yini, 2019). The robustness is measure by determining the ability of the steganographic methods to resist from any transformation such as compression or decompression of stego text.

Table 3 show the embedding capacity of hidden message using various size of hidden message with cover text size of 750 characters.

Table 3: Embedding Capacity of Hidden Message

Hidden Message Size (character)	Capacity
26	3.46%
28	3.73%
34	4.50%
40	5.33%
43	5.70%
67	8.90%
198	26.40%
255	34.00%
267	Fail

Table 3 show that the maximum capacity ratio using this technique is 34% with the

November-December 2019
ISSN: 0193-4120 Page No. 6030 - 6037

maximum of hidden message can be embed into the cover text is 255 characters. This is because the maximum RGB color value is 255,255,255 which mean that the maximum value of z is 255. The table also shown that, the proposed technique fail to embed the hidden message with the size more than 255 characters.

6. Conclusion

This mapping technique successfully conceals a hidden message into cover text by converting an ASCII character of hidden message into 3D representation. The 3D representation has been mapping with RGB color to format a selected character in cover text. PRNG has been used to find a random location for hiding purpose. Although the hidden message is fully conceal in cover text with a high capacity but the representation of character A to Z was form a similar pattern of *x* and *y* value. The Jaro-Winkler score equal to 1 show the similarity of cover text and stego text. The imitation of this study is the *z* value is always in a sequential which tend to steganalysis to extract the hidden message. The future research can be look on how to represent a hidden message with a different representation to avoid a similar pattern.

References

[1] Ahvanooey, M. T., Li, Q., Hou, J., Rajput, A. R., & Yini, C. (2019). Modern Text Hiding , Text Steganalysis , and Applications : A Comparative Analysis. *Entropy*, *21*(4), 1–29. https://doi.org/10.3390/e21040355

[2] Ahvanooey, M. T., Li, Q., Shim, H. J., & Huang, Y. (2018). A Comparative Analysis of Information Hiding Techniques for Copyright A Comparative Analysis of Information Hiding Techniques for Copyright Protection of Text Documents. *Security and Communication Networks*, (April). https://doi.org/10.1155/2018/5325040

[3] Al-Asadi, S. A., & Bhaya, W. (2016). Text Steganography in Excel Documents Using Color and Type of Fonts. *Research Journal of Applied Sciences*, *11*(10), 1054–1059.

[4] Al-Azzawi, A. F. (2018). A Multi-Layer Hybrid Text Steganography For Secret Communication Using Word. *International Journal of Network Security & Its Applications*, *10*(6), 1–12. https://doi.org/10.5121/ijnsa.2018.10601

[5] Baawi, S. S., Mokhtar, M. R., & Sulaiman, R. (2018). A Comparative Study on The Advancement of Text Steganography Techniques in Digital Media. *ARPN Journal of Engineering and Applied Sciences*, *13*(5), 1854–1863.

[6] Bhat, D., Prabhu, S., & Renuka, A. (2017). Information Hiding through Dynamic Text Steganography and Cryptography. In *International Conference on Advances in Computing, Communications and Informatics* (pp. 1826–1831).

[7] Bhattacharyya, S., Banerjee, I., & Sanyal, G. (2011). A Survey of Steganography and Steganalysis Technique in Image, Text, Audio and Video as Cover Carrier. *Journal of Global Research in Computer Science*, *2*(4), 44–54.

[8] Bhattacharyya, S., Indu, P., Dutta, S., Biswas, A., & Sanyal, G. (2011). Hiding Data in Text Through Changing in Alphabet Letter Patterns (CALP). *Journal of Global Research in Computer Science*, *2*(3), 33–39.

[9] Bhaya, W., Rahma, A. M., & Al-nasrawi, D. (2013). Text Steganography Based on Font Type in MS-Word Documents. *Journal of Computer Science*, *9*(7), 898–904. https://doi.org/10.3844/jcssp.2013.898.904

[10] Chaudhary, S., & Dave, M. (2016). Text Steganography Based on Feature Coding Method. In *International Conference on Advances in Information Communication*

November-December 2019
ISSN: 0193-4120 Page No. 6030 - 6037

Technology & Computing (pp. 5–8).

[11] Chaundhary, S., Dave, M., & Sanghi, A. (2016). Aggrandize Text Security and Hiding Data through Text Steganography. In *Conference: 2016 IEEE 7th Power India International Conference (PIICON)*.

[12] Din, R., & Utama, S. (2018). Analysis Review of Feature-Based Method in Term of Verification and Validation Performance. *Journal of Telecommunication, Electronic and Computer Engineering*, *10*(2–4), 173–177.

[13] Elmahi, M. Y., Wahbi, T. M., & Sayed, M. H. (2017). Text Steganography Using Compression and Random Number Generators. *International Journal of Computer Applications Technology and Research*, *6*(6), 259–263.

[14] Fridrich, J. (2010). *Steganography in Digital Media: Principles, Algorithms, and Applications* (First Edit). New York: Cambridge University Press.

[15] Hmood, A. K., Jalab, H. A., Kasirun, Z. M., Zaidan, B. B., & Zaidan, A. A. (2010). On the Capacity and Security of Steganography Approaches-An Overview.pdf. *Journal of Applied Sciences*, *10*(16), 1825–1833.

[16] Ingemar, J. C., Matthew, L. M., Jeffrey, A. B., Fridrich, J., & Kalker, T. (2008). *Digital Watermarking and Steganography* (Second). Burlington, USA: Morgan Kaufmann Publishers.

[17] Iyer, S. S. (2017). Practical Evaluation and Comparative Study of Text Steganography Algorithms. *International Journal of Advance Engineering and Research Development*, *3*(4), 277–283.

[18] Khairullah, M. (2019). A Novel Steganography Method using Transliteration of Bengali Text. *Journal of King Saud University - Computer and Information Sciences*, *31*(3), 348–366. https://doi.org/10.1016/j.jksuci.2018.01.008

[19] Koley, S., & Mandal, K. K. (2017). Number System Oriented Text Steganography in Various Language for Short Messages. In J. K. Mandal, P. Dutta, & S. Mukhopadhyay (Eds.), *Computational Intelligence, Communications, and Business Analytics: International Conference* (1st ed., Vol. 1, pp. 552–566). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-10-6427-2

[20] Krishnan, R. B., Thandra, P. K., & Baba, M. S. (2017). An overview of text steganography. In *2017 4th International Conference on Signal Processing, Communication and Networking, ICSCN 2017* (pp. 0–5). https://doi.org/10.1109/ICSCN.2017.8085643

[21] Kumar, P., & Sharma, V. K. (2014). Information Security Based on Steganography & Cryptography Techniques : A Review. *International Jornal of Advanced Research in Computer Science and Software Engineering*, *4*(10).

[22] Lwin, T., & Phyo, S. W. (2014). Information Hiding System using Text and Image Steganography. *International Journal of Scientific Engineering and Technology Research*, *3*(10), 1972–1977.

[23] Malik, A., Sikka, G., & Verma, H. K. (2017). A High Capacity Text Steganography Scheme Based on Huffman Compression and Color Coding. *Journal of Information and Optimization Sciences*,*38*(5), 647–664. https://doi.org/10.1080/02522667.2016.1197572

[24] Osman, B., Yasin, A., & Omar, M. N. (2016). An Analysis of Alphabet-based Techniques in Text Steganography. *Journal of Telecommunication, Electronic and Computer Engineering*, *8*(10), 109–115.

[25] Saraswathi, V., & Kingskin, S. (2014). Different Approaches to Text Steganography: A Comparison. *International Journal of Research in Management and Technology*, *9359*(11), 124–127.

[26] Sharma, S., Gupta, A., Trivedi, M. C., & Yadav, V. K. (2016). Analysis of Different Text Steganography Techniques : A Survey.

In *2016 Second International Conference on Computational Intelligence & Communication Technology* (pp. 130–133). https://doi.org/10.1109/CICT.2016.34

[27] Shivani, Yadav, V. K., & Batham, S. (2015). A Novel Approach of Bulk Data Hiding using Text Steganography. *Procedia Computer Science*, *57*, 1401–1410. https://doi.org/10.1016/j.procs.2015.07.457

[28] Singh, H., & Diwakar, A. (2014). An Intelligent Approach for Secure Data Transmission using Text Steganography. In *International Congress on Computer, Electronics, Electrical, and Coomunication Engineering* (Vol. 59, pp. 13–17). https://doi.org/10.7763/IPCSIT.2014.V59.3

[29] Singh, S., & Singh, A. (2013). A Review on the Various Recent Steganography Techniques. *International Journal of Computer Science and Network*, *2*(6), 142–156.

[30] Sloan, T., & Hernandez-castro, J. (2015). Forensic Analysis of Video Steganography Tools PrePrints PrePrints. *PeerJ Computer Science*, (May), 1–14.

[31] Srikanth, P., Mehta, A., Yadav, N., Singh, S., & Singhal, S. (2017). Encryption and Decryption Using Genetic Algorithm Operations and Pseudorandom Number. *International Journal of Computer Science and Network*, *6*(3), 455–459.

[32] Stojanov, I., Mileva, A., & Stojanovi, I. (2014). A New Property Coding in Text Steganography of Microsoft Word Documents A New Property Coding in Text Steganography of Microsoft Word Documents. In *SECURWARE 2014 : The Eighth International Conference on Emerging Security Information, Systems and Technologies*.

[33] Sumathi, C. P., Santanam, T., & Umamaheswari, G. (2013). A Study of Various Steganographic Techniques Used for Information Hiding. *International Journal of Computer Science & Engineering Survey*, *4*(6), 9–25. https://doi.org/10.5121/ijcses.2013.4602

[34] Tang, X., & Chen, M. (2013). Design And Implementation Of Information Hiding System Based On RGB. In *IEEE-3rd International Conference on Consumer Electronics, Communications and Networks* (pp. 217–220).

[35] Wang, X., & Li, H. (2014). Research on Information Hiding Method Based on Word Text. *Advanced Materials Research*, *930*, 2815–2818. https://doi.org/10.4028/www.scientific.net/AMR.926-930.2815

[36] Zielińska, E., Mazurczyk, W., & Szczypiorski, K. (2014, March). Trends in Steganography. *Communications of the ACM*, *57*(3), 86–95. https://doi.org/10.1145/2566590.2566610