Bulletin of Electrical Engineering and Informatics
Vol. 8, No. 1, March 2019, pp. 291~297
ISSN: 2302-9285, DOI: 10.11591/eei.v8i1.1440

The evaluation performance of letter-based technique on text steganography system

Roshidi Din[1], Rosmadi Bakar[2], Sunariya Utama[3], Jamaluddin Jasmis[4], Shamsul Jamel Elias[5]
[1,2,3]School of Computing, College Arts and Sciences, Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia
[4]UiTM (Melaka) Jasin Campus, 77300 Merlimau, Melaka, Malaysia
[5]Universiti Teknologi MARA (UiTM) Kedah, Malaysia

Article Info

Article history:

Received Sep 30, 2018
Revised Nov 2, 2018
Accepted Jan 10, 2019

Keywords:

Capacity of size bit
Embedding time
Feature-based
Logical design

ABSTRACT

Steganography is a part of information hiding in covering the hidden message in any medium such as text, image, audio, video and others. This paper concerns about the implementation of steganography in text domain called text steganography. It intends to concentrate on letter-based technique as one of the representative techniques in text steganography. This paper displays some techniques of letter-based that is integrated in one system technique displayeed in a logical and physical design. The integrated system is evaluated using some parameter that is used in order to discover the performance in term of capacity after embedding process and the time consuming in the development process. This paper is anticipated to contribute in describing the implementation of the techniques in one system and to display the performance some parameter evaluation.

Corresponding Author:

Roshidi Din,
School of Computing, College Arts and Sciences,
Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia
Email: roshidi@uum.edu.my

1. INTRODUCTION

Steganography is best known as the knowledge of art and science in hiding messages of data that cannot be recognized by human vision. The private information is a main idea of steganography in developing the performance implemented as a part of classification in information hiding and information security [1-3]. The category of steganography is divided into two kind part implementations that are technical steganography and natural language steganography. Steganography that is developed via text system could be called natural language steganography, whereas steganography in other medium is called technical steganography [4]. Text steganography able to conceal the hidden information become unnoticeable [5] by a stranger to detect as it is directed to send the hidden information to the true receiver [6]. There are two other sub-categories of natural language steganography. The first is linguistic steganography. This type of steganography is dependable with linguistic order of sentence in the text. The second is text steganography that manipulates parts of text that are word, line, space and any other criteria related with the text in order to hide the message [7, 8]. This paper intends to focus on text steganography category by looking at feature-based technique. This technique covers hidden message based on pattern letter or length of the word that conceals and seem nothing changes happen in the text. Meanwhile, feature-based can define as a technique that altered unique feature characteristic in text based on code word [9]. It covers hidden message based on pattern or length letter of the word that conceals and seem nothing changes happen in the text [10].

One of feature-based technique is letter-based that embed the technique the text in any language that use A-Z letters. This this paper compare the three techniques on letter-based which are Changing Alphabet Letter Patterns (CALP) [11] Curve in Character Subheading (CURVES) and Vertical Straight Line (VERT) [12]. CALP uses the unused symbol of ASCII number system for the pattern. The Mapping sequences of this

scheme focus on letters '*i'* and *'j'* because of those letter have dot (.) in the letter for embedding 0 bit. In embedding 1 bit *'a', 'A',* and '*c'* character will be used. Secondly, CURVE is a scheme that divides the English letters into two groups based on the outline of letter shape. The CURVE scheme as stego key is based on letter which contain full or partial curvature, which mean that every word that has a curve will be identified as group hide 0 bit. In contrast, a letter without any sort of curvature is identified as well as a group 1 bit. Thirdly, VERT is a scheme that divides English letter into two categories but it is divided by form of straight line in of the letters. Similar with CURVES, VERT technique is divides the group letter into two group names. The letter that does not have a vertical straight line hide 0 bit and the letter that has vertical line character hides 1 bit.

2. SYSTEM DEVELOPMENT

These three techniques are designed in one interface system in order to compare of each technique with some parameter. The process showing how the system works begin with prepares the cover text as the medium of text that will embed. Then, the hidden message are also prepares as the input text that will be embedded in the cover text. The hidden message will be converted into binary bits that embed in the letter chooses based on the selected technique among CALP, CURVE, VERT techniques. The system design uses in this paper will be based on logical and physical design that will be explained in the next section.

2.1. Logical model

The model of the letter-based technique begins with a logical design as a guidance to develop and implement the technique in the form of a system. The logical design is a demonstration in term of graphical system that shows the process and the flows of data into and out of the system. In this paper, a logical design used is a case model and a sequence diagram. The model of use case diagram is shown in Figure 1.

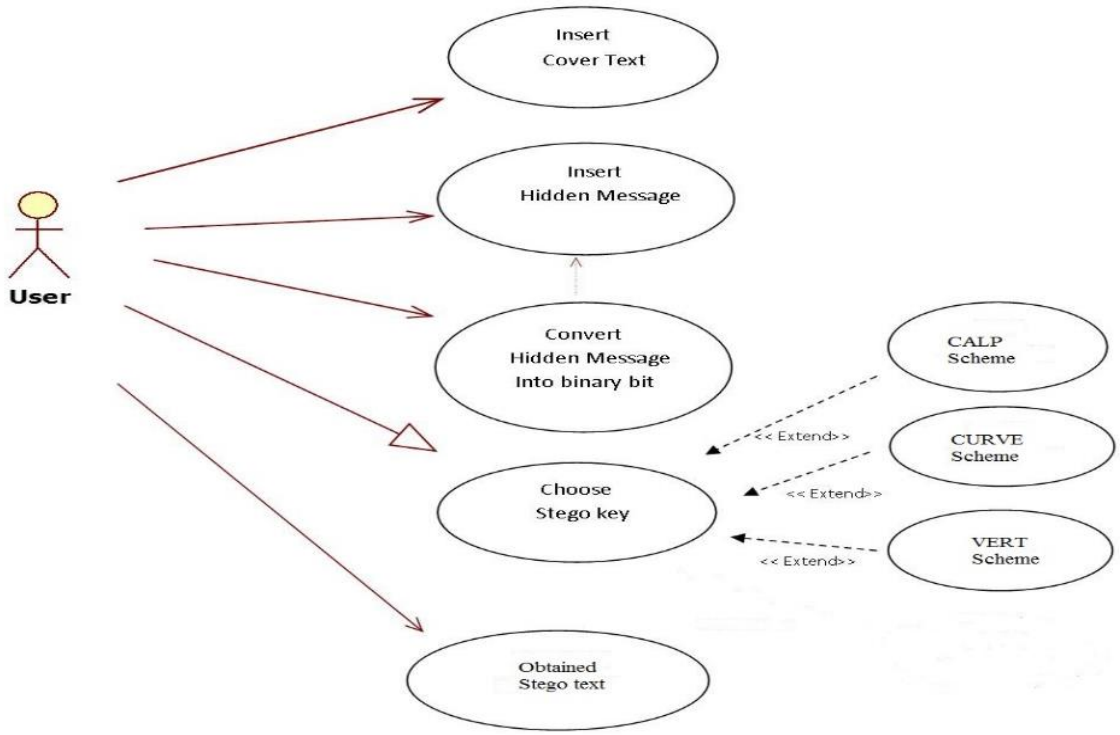

Figure 1. Use case diagram of text steganography technique

Figure 1 displays the requirement in letter-based technique of this study. It requires the process of inserting cover text and hidden message, and then converting the hidden message into binary bits. This paper also implement the sequence diagram to show the process of the system. Sequence diagram is a graphical order to presents the arrangement operation process in the system. The sequence diagram of the system is shown in Figure 2 as follow.

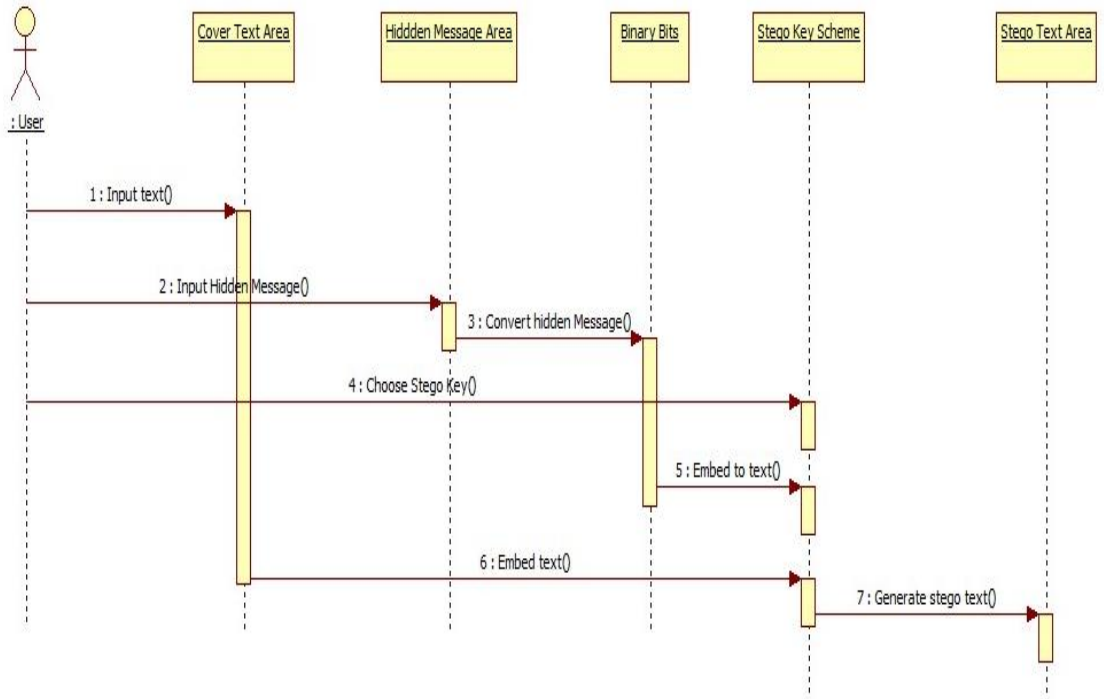

Figure 2. Chart of sequence diagram of text steganography system

Figure 2 shows the order process of the proposed system in the cover text and the hidden message areas. The hidden message is converted into binary bits. Then, the user has chosen the stego key of the selected technique that could generate the stego text in text area.

2.2. Logical model

The physical design is a graphical representation of a system that showing internal and external entities of the system as well as the data flows of these entities. In this section the interface of model letter-based usingCALP, CURVE and VERT schemes are shown in Figure 3 as follows.

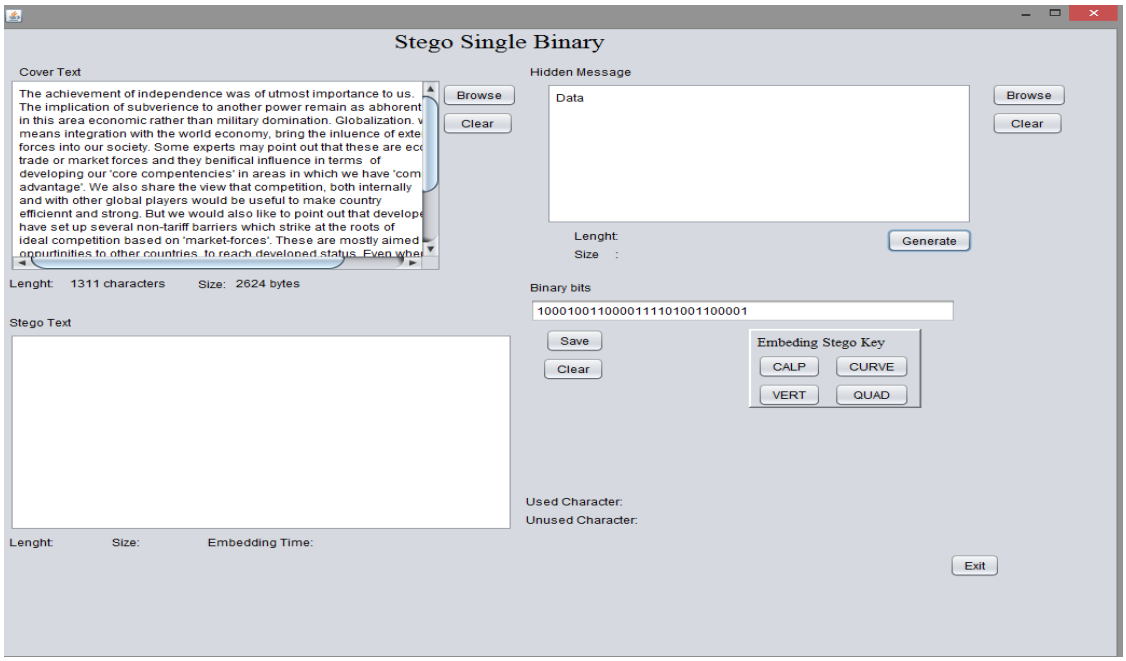

Figure 3. Physical design of the interface letter-based system

3. EVALUATION RESULT

In order to fulfill an essential and specific requirement of methods, it is important to do the evaluation procedures. The function of evaluation is to predict the quality of requirements that will be used in developing the system [13-16]. This paper use two parameters in order to measure these three techniques which are capacity of size bit and embedding time of the systems.

3.1. Capacity of size Bit

The input datasets in the cover text are investigated using five different types of length character as following respectively. The size bits are based on the capacity of size bit of the datasets. In Table 1 illustrates the capacity size bit of a cover text with the hidden message. It is similar to the experiment of length character that provides five cover texts with hidden messages.

Table 1. The size Bit of cover text and hidden message

No.	Size bit of Cover Text (CT) (Kb)	Size bit of Hidden Message (HM) (Kb)
1	2508.88	0.126
2	2742.14	0.202
3	2099.13	0.280
4	2643.24	0.356
5	2777.29	0.430
6	-	0.504
7	-	0.576
8	-	0.650
9	-	0.726
10	-	0.804

Based on Table 1, there are five random capacity of size bit of cover text and 10 hidden messages from the lowest to highest capacity of the size bit. Figure 4 show the amount of size bit of stego text using CALP technique as follows.

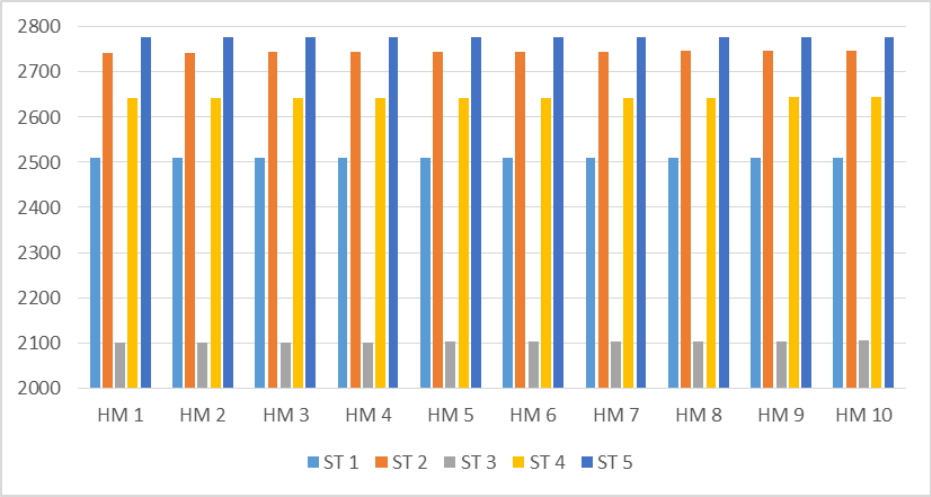

Figure 4. The result of capacity of size bit on CALP technique

In Figure 4 shows the performance of CALP in parameter of the capacity of size bit on stego text toward the hidden message. The larger of the size bit of the hidden message is the larger of size bit capacity of stego text. For example, stego text 1 (ST1) begins with 2509.58 kb until it reaches 2516.74 kb in the last hidden message. As for size bit CURVE technique it is shown in Figure 5 as follow.

Based on Figure 5, the size bit of stego text CURVE also increases if the size of the hidden message is larger. It shows how ST1 in hidden message 1 is similar to the size bit 2509.28kb and in the last hidden message 10 when it reaches 2513.36kb. However, the size bit CURVE is less than CALP technique's size bit.

For example, ST 1 in CURVE technique only started with 2509.28 kb until 2513.36 while the CALP technique has the highest size bit reach 2516.74 kb. Finally, the capacity of size bit of VERT technique is shown in Figure 6 as follow.

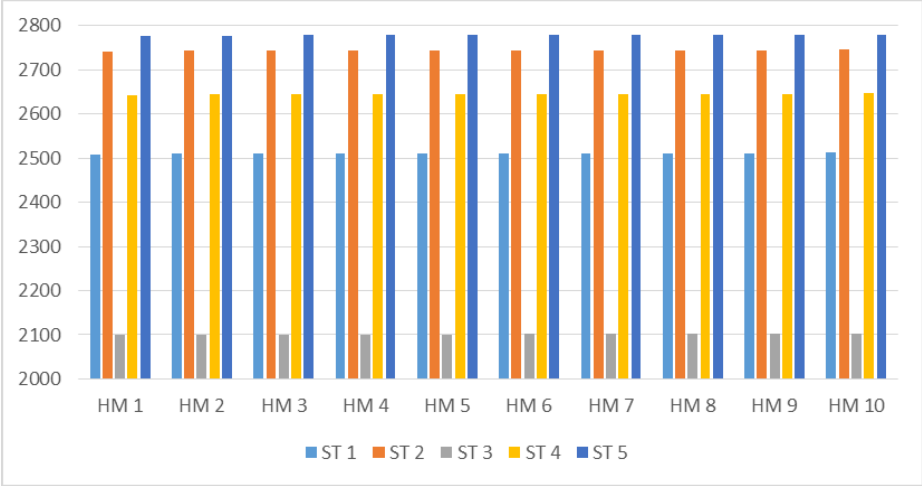

Figure 5. The result of capacity of size bit on CURVE technique

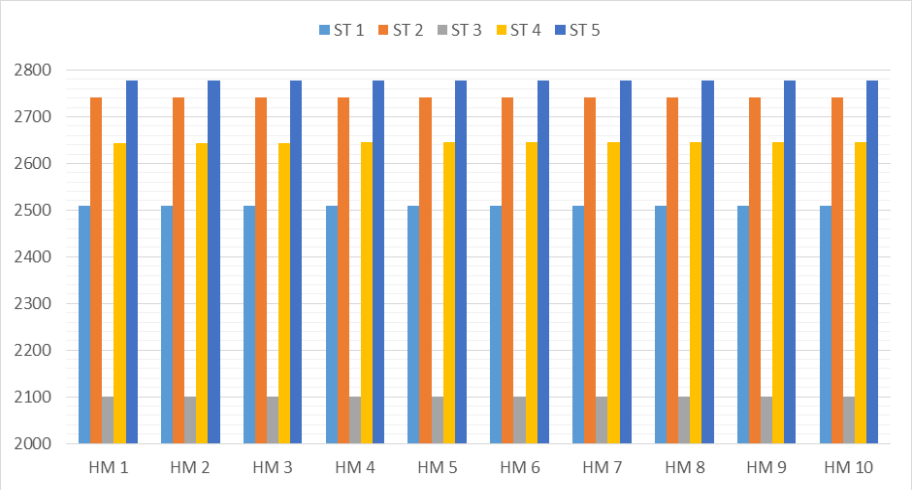

Figure 6. The result of capacity of size bit on VERT technique

In Figure 6 shows that the VERT technique also increased in the number of the size bit capacity, however, it is not significant like other technique. A hidden message 1 begins with 2508.97kb and the last hidden message only reaches 2509.85kb. Based on the comparison of these three figures, the number of size bit of stego text using VERT technique has less number of capacities compared to CURVE and CALP techniques.

3.2. Embedding time

This paper uses parameter of embedding time in order to discover an execution time of each technique toward different size bit of cover text. Figure 7 shows the embedding time of the three stego texts (ST1, ST2 and ST3) in CALP technique.

Figure 7 has illustrated the result of embedding time on CALP technique. It found that the longest time of the embedding process is occurred in stego text ST2 because cover text 2 has the highest size bit capacity which is 2742.15 kb. However, the first hidden message HM 1 is the longest time for embedding process, while the second hidden message HM2 is the longest of the execution time is occurred in stego text ST2. Next, the result of embedding time in CURVE technique is shown in Figure 8 as follows.

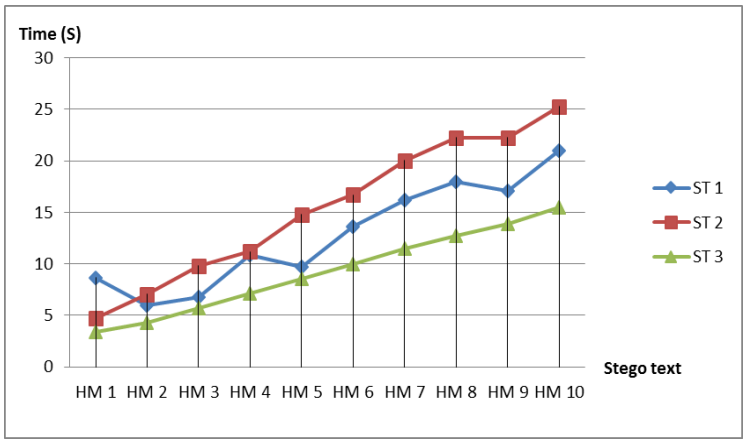

Figure 7. The result of embedding time on CALP technique

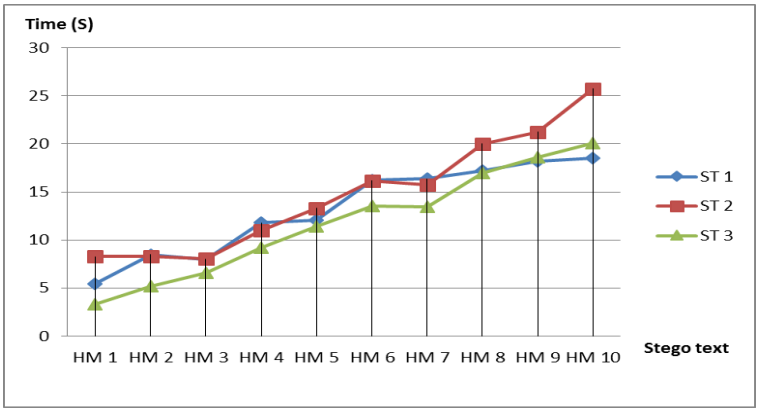

Figure 8. The result of embedding time on CURVE technique

The embedding time process among three stego texts such as ST1, ST2 and ST3 in CURVE technique become slightly random in the longest time consuming. As shown in Figure 8, it is found that between hidden message HM2 and hidden message HM7 that used in the stego text ST1 and stego text ST2 have almost produce the similar execution time in the embedding process. However, the longest time in embedding process is produced by stego text ST2. Moreover, even the stego text ST3 is the fastest embedding time process, hidden message HM10 has produced the fastest embedding time for stego text ST1 through CURVE technique. Then, The result of embedding time on VERT technique is shown in Figure 9.

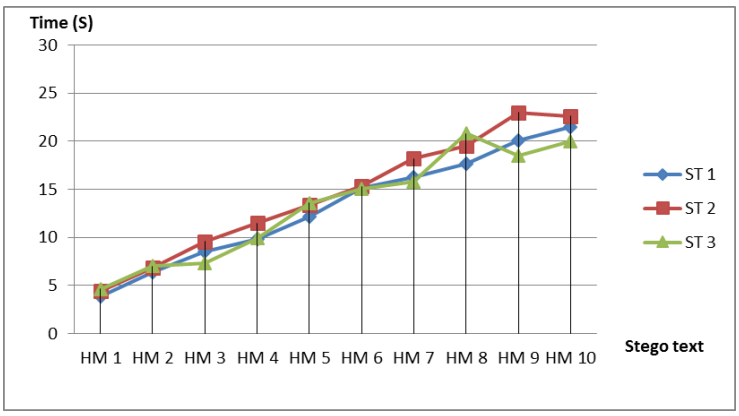

Figure 9. The result of embedding time on VERT technique

Figure 9 shows the embedding process of VERT technique that almost have similar execution time for three stego texts such as ST1, ST2 and ST3 . In general, the VERT can be assumed have the fastest execution time of embedding process compared to all techniques. It is because the execution time of embedding processes of all hidden message are not reach 25 seconds.

4. CONCLUSION

This paper introduces the letter-based technique of text steganography as the part of category of feature-based technique. It is comparing three techniques in one system that consist with CALP, CURVE and VERT techniques. These three techniques is integrated in one interface system which are displayed in use case and sequence diagram of logical design and physical design. This paper also shows the evaluation of these three techniques in the system through capacity of size bit and embedding time. Capacity of size bit is discovered the highest hidden message that embeds in cover text which is generating the highest capacity of size bit. Generally, it can be concluded that the highest capacity of size bit is produced through CALP technique and the lowest capacity of size bit is produced through VERT technique. Besides, it also can be concluded that the higher size bit of hidden message, the longer time is needed in the embedding process.

ACKNOWLEDGEMENTS

The author wish to thank the Ministry of Higher Education Malaysia in funding grant under the University Grant Non-PI with S/O 13850, RIMC of Universiti Utara Malaysia, Kedah.

REFERENCES

[1] Iyer S., Laktharia. S. S. L. New Robust and Secure Alphabet Pairing Text Steganography Algorithm. *International Journal of Current Trends in Engineering & Research (IJCTER)*, 2016; vol. 2 no. 7 pp. 15-21.
[2] Nasab. M.V., Shafiei B. M. Steganography in programming. *Australian Journal of Basic and Applied Sciences*; 2011; Vol. 5 no.12, pp. 1496-1499.
[3] Changder, S. Debnath N. C., Ghosh D. *A New Approach to Hindi Text Steganography by Shifting Matra*. International Conference on Advances in Recent Technologies in Communication and Computing, 2009; pp, 199-202.
[4] Huijuan Yang and A. C. Kot, "Text document authentication by integrating inter character and word spaces watermarking," *2004 IEEE International Conference on Multimedia and Expo (ICME) (IEEE Cat. No.04TH8763)*, Taipei, 2004, pp. 955-958 Vol.2.
[5] Sing, H. Sing, P. K. Saroha, K. *A Survey on Text based Steganography*. Proceeding of 3rd National Conference; INDIACom-2009 Computing For Nation Development, 2009.
[6] Nasab, M. V., Shafiei, B. M. Steganography in Programming. *Australian Journal of Basic and Applied Sciences*; 2011; vol.5 no. 12, pp. 1496-1499.
[7] Roy, S. Manasmiti, M. *A Novel Approach to Format Based Steganography*. ICCCS'11, 2011; pp. 511-517.
[8] Din, R. Samsudin, A. Muda, T.Z.T. Amphawan, A. Omar, M. N. Fitness Value based on Evolution Algorithm Approach for Text Steganalysis Model. *International Journal of Mathematic Models and Methods in Applied Science*. 2013; vol 7 no. 5: 551-558.
[9] Agarwal, M. Text Steganography Approaches: A Comparison. *International Journal of Network Security & Its Applications (IJNSA)*, 2013; vol.5 no.1, 91-106.
[10] Changder, S. Debnath, N. C. Ghosh, D. *A New Approach to Hindi Text Steganography by Shifting Matra*. 2009 International Conference on Advances in Recent Technologies in Communication and Computing: 2009; pp. 199-202.
[11] Bhattacharya, S. Indu, P. Duta, S. Biswas, A. and Sanyal, G. Hiding Data in Text Through in Alphabet Letter Patterns (CALP). *Journal of Global Research in Computer Science*, 2011; vol. 2 no.3 pp. 33-39.
[12] Dulera, S. Jinwala, D. Dasgupta, A. Experimenting with the Novel Approaches in Text Steganography. *International Journal of Network Security & Its Applications (IJNSA)*, 2011; vol.3 no. 6, pp. 213-225.
[13] Dasso, A. Funes, A. A. Verification, Validation, and Testing in Software Engineering. London: Idea Group Publishing, 2005.
[14] Catal, C. Performance Evaluation Metrics for Software Fault Prediction Studies. *Acta Polytechnica Hunga*rica, 2012; vol.9 no.4, 193-206.
[15] Suhartono et al, Speaker Recognition in Content-based Image Retrieval for a High Degree of Accuracy, *Bulletin of Electrical Engineering and Informatics (BEEI)*, September 2018, vol. 7, No. 3, pp. 350-358.
[16] Joseph A. Issa., Performance Evaluation and Estimation Model Using Regression Method for Hadoop Word Count". *IEEE Access*. 2015; vol. 3: pp.2784-2793. DOI: 10.1109/ACCESS.2015.2509598