

TextStegoHYB-Secured Way of Information Hiding

Mr. Yogesh N. Chaudhari¹, Dr. Bhojraj H. Barhate²

¹Assistant Professor, Department of Computer, KCES's Institute of Management and Research, Jalgaon.

Email: yogeshchaudhari@imr.ac.in, chaudhariyogn@gmail.com.

ORCID ID: 0009-0001-9304-164X

²Vice-Principal, Department of Computer, Bhusawal Art, Science & P. O. Nahata Commerce College, Bhusawal.

Email: bhbrama123@gmail.com

Date of Submission: 12th October 2023 Revised: 16th November 2023 Accepted: 29th November 2023

How to Cite: Mr. Yogesh N. Chaudhari et al., (2023), *TextStegoHYB-Secured Way of Information Hiding*, *International Journal of Applied Engineering and Technology* 5(4), pp. 932-938.

ABSTRACT: In Today's world importance of information technology is rapidly increases. Information hiding is also a one kind of technology that has major demands in the field of Information Technology. That is because Human beings are continually trying to learn new things that were hidden and beyond his knowledge, also those things which are confidential and Secret. Even the least curious people are interested in getting knowledge about the secrets of others. However, maintaining the secrecy of important documents from other influences is the major task today. That's the reason why researchers take much interest in the field of steganography. As information is becoming the most valuable property in the current era, the process of encoding messages known as encryption is playing a massive role today. Steganography, the art of invisible communication, will provide the solution of locking the information. This is the art of secret science. In this paper, we propose a new approach of information hiding called TextStegoHYB, which provides two-way security while the transmission of short text messages through the short messaging services to end users. The TextStegoHYB technique is a more trusted technique because we have assessed the efficiency of the proposed technique in terms of Payload capacity, Imperceptibility, Security and Robustness. The experiments study confirm that the TextStegoHYB can prevent message from an attacker's attention by its Unique's embedding techniques. Additionally, we have compared the results of experimental study with the existing techniques to show the advantage of the proposed technique. As per the best of our knowledge, this is the new approach of text steganography technique that provides peer-to-peer secure mechanism for communication of the short message using numerical encoding.

Keywords: Information Hiding, Text steganography, Secret Communication, SMS.

I. INTRODUCTION

The increasing use of the Internet and smartphones has a strong social and commercial impact on daily life. These new technologies benefit people around the world and make it possible to store, process and retrieve information conveniently and widely accessible. As text messaging has become a popular and easy form of communication, there are concerns about data leakage attacks such as hacking, hijacking, and phishing. Users such as investigators, journalists, judges, and election officials communicate with each other via short message services (SMS) or social media applications[1], [2], [3]. Smartphone users often send sensitive information such as bank account details (for example, account information, passwords, and transaction information), secret missions, confidential meetings, and private identities of family members via SMS or text messages using small text messaging applications. However, surprisingly, standard SMS services and social networks do not guarantee the security of this type of digital data sent over the network. In this case, it is necessary to ensure secure communication between smartphone users. Because text messaging via SMS is so common on smartphones. Functionally, text messages are sent unencrypted between smartphone users and service providers (SMS service centers, social media service providers, etc.) over an available network. The content of text messages is stored by service providers on easily accessible servers. Many techniques have been proposed in the last two decades, but these techniques have been some limitations in terms of Embedding Capacity, Robustness, Imperceptibility, and distortion attacks, in other words, you cannot embed a large amount of secret information in a short message.[4], [5]

The major contributions of these articles are:

- ❖ **First.** In Section II, Background Study, Introduces some basic and related information papers on text steganography.
- ❖ **Second,** In Section III, we propose a new steganography text called TextStegoHYB, which guarantees safety Message when sent via SMS. TextStegoHYB protects confidential information Generate a hidden message (HM) containing the following confidential information.
- ❖ **Third,** In Section IV, we develop an application based on TextStegoHYB and conduct practical experiments to evaluate it, using fifteen social networking and instant messaging applications such as: communication channels. Examples are provided for analysis and compare the proposed technique with the existing one technique based on common criteria.
- ❖ **Finally,** In Section V, we conclude the article with a summary of research contribution and future work.

II. BACKGROUND STUDY

This section Firstly, gives a short description of Information Security Systems, and the Attributes of Steganography Systems, and then presented the various studies and related works on text steganography methods and presented their comparison.

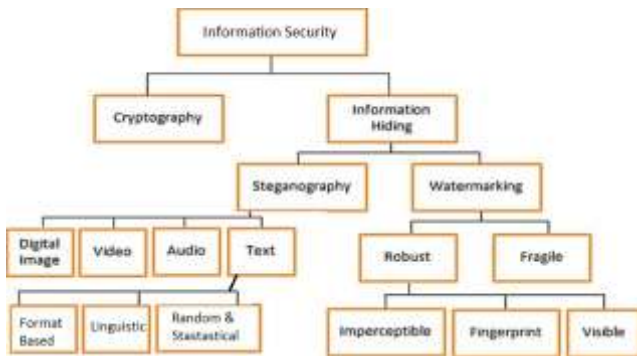


Figure 1: Information Security Systems[6]

A) Information Security

Information security is an important topic as part of user data protection. Although there are several approaches that are robust and safe. See fig (1). The security of these approaches continues to improve safer in terms of performance parameters. Generally, there are two classes Information Security Systems: Information Hiding and Cryptography [1]. The two classes can protect information; However, both approaches are different. Researcher formulated various types of data encryption techniques in [2]. Figure 1 shows the general classification of the data security mechanism that combines the three techniques presented: Steganography, Watermarking and Cryptography. Steganography is divided into: Image, Audio, Video and Text as per the

cover media used and watermarking are divided into Robust and Fragile Watermark.[6]

Generally, cryptography added a plain text into cipher text, it should be reversible without data loss. The aim of cryptography is to avoid unauthorized access to the Secret Message (SM) by crawling the original form of its content. On the other hand, Steganography is a technique that hides an SM in cover of an innocent media (text, image, video, and audio) for the purpose of secure transmission[7], [8].

Table 1. Comparison among different information security techniques.[6]

Characteristics	Steganography	Watermarking	Cryptography
Goal	Protects secret data	Protects legitimacy	disorganizes the
Cover Choosing	free cover choice	Limit	Not use
Challenges	Imperceptibility,	Robustness	Robustness
Keys	Possible	Possible	Necessary
Output	Stego-	Watermark	Encryption
Visibility	Definitely	Occasional	Constantly
The system is invalid if	Noticed	Detached or	Decryption
Attacks	Steganalysis	image	Cryptanal

B) Attributes of Steganography Systems

There are many attributes that are considered while evaluations of any text steganography techniques, however some common parameters play important role in determining the efficiency of the text steganography techniques. Invisibility, Embedding Capacity, Robustness, Security are the five attributes of steganography forced to hide secret data.

1) Embedding Capacity (EC):

Because the Encoding algorithm converts each letter SM into two part and generate the word to generate CM and then that SM and CM is combined to generate the Stego-Object to send on the communication channel, important aspect of the Steganography system to maximize the embedding capacity without noticeable change because the change may attract the attacker’s attention. We have suggested the proposed techniques for SMS. But it will embed the text with space symbols allowed. It uses the ZWC characters to hide the space symbols used the SM.[6]

2) Imperceptibility:

Imperceptibility is an important aspect in steganography and aims to hide secret data in other media. The human eye cannot understand this even with statistical methods [11]. Statistical methods offer attackers an advantageous

opportunity to determine whether secret data is being transmitted during communication between two parties. Therefore, major media outlets are not expected to experience significant changes in statistical standards due to the inclusion of the classified data; That is, if similar statistical data is found in the original and stego files, we can assume that the security is high enough to allow data communication. When sharing over insecure networks, the quality of external media must be maintained despite interference from the embedding process. Our proposed systems gives highest Imperceptibility because it generate the cover text and do not directly include the SM in any hidden characters that will detected in terms of space before or after the Cover Text in special kind of editors.[1], [6], [9], [10]

3) Robustness(DR):

Basically, one of the most important criteria in steganography is the reversibility of the CM. To solve this task, we does not includes any prior and trailer ZWC characters that will be deleted after change the contents of the CM by attackers. That results in a lower chance of Secret Message destruction.

Therefore, if the attacker uses the LTRIM or RTRIM function to remove any ZWC appended in terms of prior and trailer of CM, it will not affect the actually meaning of the text that's give highest robustness and can be extracted using the TextStegoHYB technique. [1], [6], [9], [10]

4) Security:

Since TextStegoHYB offers an proper balance between three other criteria - medium EC, highest Robustness, highest invisibility and Highest Security - it is able to perfectly protect against multiple attacks. Therefore, the eavesdropper (attacker) cannot decode or even recognize the existence of message. [1], [6], [9], [10]

C) Related Studies

Previously, various researchers had presented several Approaches to Ensuring the Secure Transfer of Confidential Information on communication channels using data hiding techniques in media cover. Majority of the existing literature work focuses on data hiding techniques which uses the images, videos and audio files to embed the secret information [1], [6], [9], [10]; but relatively Various techniques have been proposed to hide secrete information in the cover text. text messages contain a limited number of words and symbols It is challenging task to change text content to hide data Short text on the cover. Basically include text messages writing styles such as special phrases, shorthand acronyms and emoticons . Text data A masking technique called UniSpaCh is presented in[11]., which generates and isolates the binary string SM 2-bit classification (e.g.for example "10, 01, 00 and 11"). That too replaces every 2

bits with a special space (e.g. Thin, Hair, Six Per-Em and punctuation). Finally, integrate the generated file spaces in special places, e.g. between words, between sentences, at the end of a line and between paragraphs. However, this approach guarantees a high level of invisibility because of the cover text but has a low capacity (two bits per space) and does not applied for the application which integrates long secret message into a short cover message. New A text steganography algorithm called AH4S was introduced in [12], which deals with the structure of the Omega network SM cache in the generated CM. Select the letter z SM and uses the Omega network to create two connected letters based on the selected letter and searches dictionary where you can hide the word from the corresponding English covert text. In which it produced two letters and repeated the same process all letters SM. For example, to hide "A", generate long, unknown text like CM from "A" and then this inserts two spaces into the generated cover text to hide the positions of the two letters generated above CM. In practice, a long and unknown text is created for a short text SM and also creates doubt among readers. Some Effective text marking technique (TWSM) for embed secret data in short Latin texts were proposed in [10], which uses Unicode homoglyphs and special rooms to integrate Latin Secrecies Based on CM. Judging from the experimental results, this might actually be the case found that this technique enabled optimal Capacity, high invisibility and low robustness to Distortion Attacks. Smartphone users sometimes uses emoticons in everyday conversations instead of typing their feelings. Several researchers have used emoticons for this purpose Hide SM via SMS. For example, , and [13], [14] generate random text containing multiple words, e.g. B is a CM and further convert all SM characters for emoticons based on a previously defined schema (e.g.G, A="angry", B="sad", C="happy", etc.) and then integrate emoticons among the words in CM.

III. PROPOSED SYSTEMS

As Shown in Fig 2. the working of proposed techniques which embed a Secret Message (SM) in a Cover Message (CM). A stego object is produced from embedding Secret Message (SM) in Cover Message(CM).. The function that is used to hide secret messages in cover messages is called encoding algorithm. The decoding procedure retrieves the hidden secret message from stego object. We have presented the proposed techniques with both Encoding and Decoding algorithm[15]

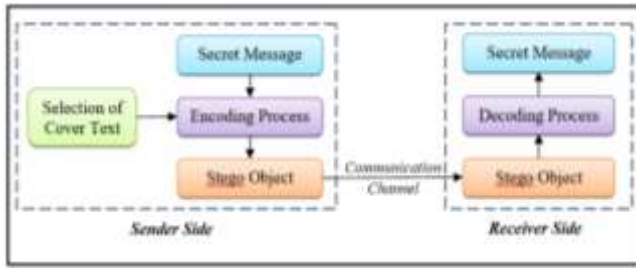


Figure 2: Working of Text Steganography System[15]

Terminology used in steganography techniques are[16]:

- Cover Message (CM):** Text, audio, video, image used for embedding data is known as Cover Message. [14], [15]
- Secret Message (SM):** The data which is to be embedded in a cover message is known as secret message. [14], [15]
- Encoding Algorithm (EA):** it hides secret message in cover message is called encoding procedure. [14], [15]
- Stego Object:** It is the resultant output obtained after embedding Secret Message and Cover Message which is known as stego object. [14], [15]
- Sender Side (SS):** This is sender side from Secret Message is sent as a SMS or others way on Communication channel. [14], [15]
- Receiver Side (RS):** This is another side which decodes the secret information from CM received on Communication channel. [14], [15]

Important requirements of a steganography method are imperceptibility, robustness, and capacity. Therefore, the primary goals of steganographic approaches are to provide higher security which make the message unpredictable. [17]

Table 2 is table of letter frequency (FOL) and number assignment as per the English dictionary [6] [15]

Letter	Frequency of letter in Percent	Number assigned	Letter	Frequency of letter in Percent	Number assigned
E	11.16	15	M	3.01	7
A	8.50	14	H	3.00	7
R	7.58	13	G	2.47	6
I	7.54	13	B	2.07	5
O	7.16	12	F	1.81	4
T	6.95	11	Y	1.78	4
N	6.65	11	W	1.29	3
S	5.74	10	K	1.10	3
L	5.49	10	V	1.01	3
C	4.54	9	X	0.29	2
U	3.63	8	Z	0.27	2
D	3.38	8	J	0.20	1
P	3.17	7	Q	0.20	0

Figure 3: Table of Letter frequency and number assignments[18]

A) Encoding Algorithm

Algorithm 1: Pseudocode of Encoding Algorithm (Sender Side Working)

Input: a Secret message (SM)

Output: StegoObject Including Cover Message (CM) + Secret Message

```

1. Secret Message(SM)->;
2. B8C Binary 8 bit combined strip;
3. B4F BIN First 4 bit Part;
4. B4S BIN Second 4 bit Part;
5. ASC ASCII;
6. arr array;
7. j = 0;
8. while i <= SM do
9.   if(SPACE)
10.    arr[j] = ADD(ZWC);
11.   else
12.    ASC = ASCII_OF(SM[i]);
13.    B8C = GET_BIN(AS);
14.    B8C = FLIP(B8C);
15.    B4F = B8C[0-3];
16.    B4S = B8C[4-7];
17.    arr[j] = GET_LOF_L(GET_HEX(B4F));
18.    arr[j] = GET_LOF_L(GET_HEX(B4S));
19.    inc i;
20.   end if
21. End While
22. i=0;
23. CM Cover Message;
24. STO StegoObject;
25. POS togetPos(3n-2,3n-1, 3n, n);
26. WRD Dictionary word;
27. while i <= arr do
28.   if(arr[i] is ZWC)
29.    STO = STO + ZWC;
30.   else
31.     bool = FOUND(GET_DICT_WRD_ST_N_END_WITH(arr[i], arr[j]));
32.     if bool = true then
33.      WRD = GET_DICT_WRD_ST_N_END_WITH(arr[i], arr[j]);
34.      POS = CALCPOS(i); //(3n-2,3n-1, 3n, n)
35.      STO = STO + DOUBLESPACE +(WRD + GET_CONJ_JOING_WORD(POS));
36.     else
37.      FWORD = GET_DICT_WRD_ST_WITH(arr[i]);
38.      SWORD = GET_DICT_WRD_ST_WITH(arr[i]);
39.      STO = STO + SINGLESPEACE + (FWORD + SWORD + GET_CONJ_JOING_WORD(POS));
40.     end if
41.   end if
42.   inc i;
43. End While
44. Return STO (SM+CM);

```

B) Decoding

Algorithm 2: Pseudocode of Decoding Algorithm (Receiver Side Working)

TextStegoHYB-Secured Way of Information Hiding

```

Input: a Stego Object (STO)
Output: Secret Message (SM)
1. Stego Object (STO)->;
2. ZWC Zero Width Character;
3. i=0;
4. SM Secret Message;
5. arr array;
6. DD DATADICT(num,val)
7.. WORD;
8.. POS = 1;
9. while i less than equals STO do
10. while j=i <= STO do
11. if(arr[i] is ZWC)
12. DD[POS] = SPACE;
13. else
14. WORD = WORD + (STO(J));
15. if(DOUBLESPACE AND
NOT(CONJ_WORD(WORD))) then
16. FL = GET_DICT_FL_OF_WORD(WORD);
17. DD[POS] <- FL;
18. inc POS;
19. LL <-GET_DICT_LL_OF_WORD(WORD);
20. DD[POS] = LL;
21. inc POS;
22. else if(SINGLESPACE AND
NOT(CONJ_WORD(WORD))) then
23. FL = GET_DICT_FL_OF_WORD(WORD);
24. DD[POS] = FL;
25. inc POS;
26. WORD = "";
27. end if
28. end if
29. inc j;
30. end while
31. inc i;
32. End While
33. B8C Binary 8 bit combined strip;
34. B4F BIN First 4 bit Part;
35. B4S BIN Second 4 bit Part;
36. ASC ASCII;
37. arr array;
38. FLN Frequency of letter number;
39. while i <= DD do
40. if(DD[i] is SPACE);
41. arr[j] = arr[i] + SPACE;
42. else
43. FLN = GET_LOF_N(DD[i])
44. B4F = GET_BIN(FLN);
45. FLN = GET_LOF_N(DD[i+1])
46. B4S = GET_BIN(FLN);
47. B8C = B4F + B4S;
48. B8C = FLIP(B8C);
49. ASC = ASCII_OF(BIN);
50. arr[i] = arr[i] + ASC;
51. end if
52. inc i;

```

```

53. End While
54. SM = arr;
55. Return SM;

```

Example:

A) Sending Side working

- Secrete Message need to send to receiver.

teju

- Final Message (Stego Message) to send to receiver:

Dijit is cries for corgo's deaths

B) Receiver Side Working

- Message received from sender is

Dijit is cries for corgo's deaths

- Secrete Message retrieved from Stego-text after decoding is:

Teju

IV. EVALUATIONS

Table 1 EMBEDDING ALGORITHM ANALYSIS OF TEXTSTEGOHYB AND EXISTING TECHNIQUES [1]

Algorit hm	Type of Embed ding	S M	CMhm (HM+ CM)	NEL RS	NCR ES (App x.)	Summary of embedding algorithm
TextSte goHYB	Bit-level	teju	Dijit is cries for corgo's deaths	1	33	It embeds a hidden string of SM in CM, which doesn't depend on CM
The AITSteg [1]	Bit-level	Ali	How are you?	1	3	It embeds a hidden string of SM in front of CM, which does not depend on CM.
UniSpa Ch [11]	Bit-level	Ali	How are/ you?	4	20	This technique embeds the secret bits by adding a special space beside of normal space into the CM, which each space refers to a 2-bit of SM binary (e.g., "00,10,01,11").
TWSM	Bit-level	Ali	How are/ you?	3	15	It utilizes the homoglyph letters and special spaces in order to embed the secret bits such that, some letters are replaced by similar letters with different codes for embedding 1-bit, and one special space is inserted between words for hiding 3-bit of secret bits.

TextStegoHYB-Secured Way of Information Hiding

- Kavita Saini”, doi: 10.1088/1757-899X/518/5/052003.
- [6] M. A. Majeed, R. Sulaiman, Z. Shukur, and M. K. Hasan, “mathematics A Review on Text Steganography Techniques,” 2021, doi: 10.3390/math9212829.
- [7] A. Das and H. U. Khan, “Security behaviors of smartphone users,” *Information and Computer Security*, vol. 24, no. 1, pp. 116–134, Mar. 2016, doi: 10.1108/ICS-04-2015-0018.
- [8] F. Al-Shaarani and A. Gutub, “Securing matrix counting-based secret-sharing involving crypto steganography,” *Journal of King Saud University - Computer and Information Sciences*, 2021, doi: 10.1016/j.jksuci.2021.09.009.
- [9] M. Aman, A. Khan, B. Ahmad, and S. Kouser, “A HYBRID TEXT STEGANOGRAPHY APPROACH UTILIZING UNICODE SPACE CHARACTERS AND ZERO-WIDTH CHARACTER,” 2017. [Online]. Available: <https://www.researchgate.net/publication/314449134>
- [10] S. Baawi, M. Mokhtar, and R. Sulaiman, “New text steganography technique based on a set of two-letter words,” *J Theor Appl Inf Technol*, vol. 95, pp. 6247–6255, 2017.
- [11] L. Y. Por, K. Wong, and K. O. Chee, “UniSpaCh: A text-based data hiding method using Unicode space characters,” *Journal of Systems and Software*, vol. 85, no. 5, pp. 1075–1082, May 2012, doi: 10.1016/J.JSS.2011.12.023.
- [12] A. M. Hamdan and A. Hamarsheh, “AH4S: an algorithm of text in text steganography using the structure of omega network,” *Security and Communication Networks*, vol. 9, no. 18, pp. 6004–6016, Dec. 2016, doi: 10.1002/SEC.1752.
- [13] Z. H. Wang, C. C. Chang, T. D. Kieu, and M. C. Li, “Emoticon-based text steganography in chat,” *PACIIA 2009 - 2009 2nd Asia-Pacific Conference on Computational Intelligence and Industrial Applications*, vol. 2, pp. 457–460, 2009, doi: 10.1109/PACIIA.2009.5406559.
- [14] Chaudhari Yogesh N and Barhate B.H, “Comparative Study of Various Steganography Techniques: A Literature Review,” *International Journal of Computer Research and Technology, A Peer Reviewed Journal*, vol. 7, no. 1, pp. 65–72, Jun. 2021.
- [15] Chaudhari Yogesh N and Barhate B.H, “A Secured Approach to Text Steganography based on Numerical Encoding,” *Phalanx: A Quarterly Review for Continuing Debate*, vol. 18, no. 1, pp. 106–117, Mar. 2023.
- [16] N. R. Zaynalov, O. N. Mavlonov, A. N. Muhamadiev, Q. Dusmurod, and I. R. Rahmatullaev, “UNICODE For Hiding Information In A Text Document,” *undefined*, Oct. 2020, doi: 10.1109/AICT50176.2020.9368819.
- [17] R. Kumar and H. Singh, “Recent Trends in Text Steganography with Experimental Study,” in *Handbook of Computer Networks and Cyber Security*, Cham: Springer International Publishing, 2020, pp. 849–872. doi: 10.1007/978-3-030-22277-2_34.
- [18] S. Roy and P. Venkateswaran, “A Text based Steganography Technique with Indian Root,” *Procedia Technology*, vol. 10, pp. 167–171, 2013, doi: 10.1016/j.protcy.2013.12.349.