ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Issue 1, 2017 January

Enhancing the Capacity of Text Steganography with the Aid of LZW Compression & LSB Technique

Graceton Thurai.A.J, Chandru.T, Ajithkumar.P
Department Of Electronics and Communication Engineering
Sri Ramakrishna Engineering College, Coimbatore-22.
Mr.S.Lakshmi Narayanan, Asst.Professor
Department Of Electronics and Communication Engineering
Sri Ramakrishna Engineering College, Coimbatore-22.

Abstract: This work focuses on enhancing the seemingly important parameters of stenography namely capacity and security of by utilizing the concepts of LZW compression technique and LSB based approach. The proposed model utilizes the cover message to fleece the clandestine data. This algorithm initially compresses the secret data and then attempts to hide the compressed secret data in the cover message. Experimental results demonstrate that the proposed method not only produces a comparatively high embedding capacity but also reduces the computational complications. Furthermore, the security of the proposed method is significantly improved by employing stego keys. The preeminence of the proposed method has been experimentally verified by associating with the recently established methods.

I. INTRODUCTION

With the progress of computer technology and the widespread usage of the Internet, it becomes more and more suitable for people to access and interchange all kinds of multimedia information like audio, video, and images. However, distribution of these types of information over public network (i.e., Internet) makes them vulnerable to attack. Thus, there is a need to have solution which can protect sensitive data. One such method is data hiding (information hiding) which plays an important role in information security over the internet. Data hiding is a process of embedding the secret data into the cover text by minimally modifying the elements of the cover text. Generally, data hiding involves both watermarking and steganography. A watermarking scheme alters a cover object to embed a message about the cover object (e.g., owner’s identifier). The main objective of the watermarking is to attain a high level of robustness i.e. it should be very difficult to remove a watermark without degrading the quality of the data object. Watermarking is mainly used for copyright protection, broadcast monitoring, image security and forensics. On the other hand, steganography is the art and science of hiding secret communication. A steganographic system thus embeds secret content in cover media (like text, image, audio, and video) so that its existence is not detected to an attacker.

In text based steganographic methods, text is used as a cover media for hiding the secret data. Text steganography is one of the hardest areas of data hiding, since the human eye is very susceptible to any change between the original and the modified texts (stego-texts) and it can be easily detected. There are mainly two parameters namely capacity and security to analyze the performance of any text steganographic method. Capacity refers to the amount of secret data that can be concealed in the carrier, and security relates to the ability of an attacker to figure out the concealed information. In this paper, both capacity and security issues have been considered to analyze the performance of the proposed text based steganography method. The main objective of this work is to obtain significant increment in the amount of secret data which is hidden in the cover medium and also design and uses stego keys for security improvement. In order to achieve this objective, a forward mail platform or more specifically the email ids and the cover message both are used to hide the secret data. In the proposed work, for capacity increment LZW data compression technique is used for compressing the secret data, because LZW has good compression ratio. In proposed work, LZW is directly applied to the secret data. Thus, the proposed method decreases the computational complexity and also increases the capacity of the secret data. The stego-keys are also employed. The stego keys are divided into two categories, namely constructed stego key which is used during embedding phase and previously constructed global stego key which is shared between sender and the receiver beforehand.

II. RELATED WORK

In this section, we discuss some of the well-known text based steganographic techniques for different languages. Wayner proposed a mimic function based technique. In this technique, the inverse of Huffman code is used by inserting a data stream of randomly distributed bits. For improving the performance, it makes use of Van Wijngaarden Grammar and

Context Free Grammar. Though, it provides resistance against statistical attacks yet it suffers from invalid syntax problems. Maher discussed a text based data hiding method which is known as TEXTO. This technique transformed ASCII data into English sentences. It converts the secret data into English words. Therefore, this method resembles with substitution cipher and reduces suspicion over the message. Sun proposed L–R scheme which uses the left and right components of Chinese characters. To conceal the secret data, the left and right components of characters are selected as candidates. If the secret data bit is ‘‘1”, the scheme modifies the candidate by adjusting the space between the left and right components otherwise leaves unchanged. To improve the L–R scheme in terms of hiding capacity, Wang modified the scheme by adding the up and down structure of Chinese characters as an extra candidate set. Apart from this, a reversible function is also added to obtain the original cover text after the initial hidden secret data has been extracted. Later, Wang used emotional icons (also called emoticons) in chat rooms over the Internet to hide the secret data. In this method, both the sender and the receiver design a table together that will be with them during their communication. The table consists of emoticons which are classified into several sets according to their meaning (like cry, smile, and laugh) and every emoticon belongs to one set. Each individual set is provided with a unique order number which is further used for hiding the secret data. Another method known as UniSpach uses spaces to hide the secret data. It uses Unicode space characters to embed the secret data in Microsoft Word document.

Satir and Isik discussed a LZW compression based text steganography method to hide the secret data. This method hides the secret data into the cover text.

III.EMBEDDING PHASE

Step 1: Apply LZW algorithm on the secret message.

Step 2: Convert the obtained LZW code into the binary form to generate a bit stream.

Step 3: Count the number of characters (without spaces) of cover text and extract the same number of bits from the bit stream.

Step 4: Now, the residual bit stream is divided into groups of 12 bits and each group is further partitioned into 9 bits and 3 bits subgroups which are known as G_1 and G_2 respectively. If the bits in the bit stream are not in the multiple of 12 then append the required number of zeros to make it the nearest multiple of 12.

Step 5: Calculate x, y and z as follows:

$x = (G1)/26$
$y = (G1)\mathrm{mod}26$
$z = G2$

Step 6: The values of x and y are converted into letters by employing Latin Square.

Step 7:Thus, using the T as a cover text and K_2 as the stego-key, the secret message is transmitted in the form of forward mail platform.

R/C	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
2	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
3	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
4	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
5	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
6	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
7	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
8	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
9	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
10	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
11	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
12	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
13	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
14	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
15	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
16	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
17	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
18	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
19	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
20	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
21	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
22	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
23	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
24	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
25	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
26	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

LATIN SQUARE TABLE

IV. Extraction Phase

Step 1: Get the stego-cover and extract the secret data bit information cover text .

Step 2.:Extract first two elements and convert them to numbers by employing Latin Square. Thus, G_1 and G_2 are obtained using following equations

G1=(x*26+y)

G2=z

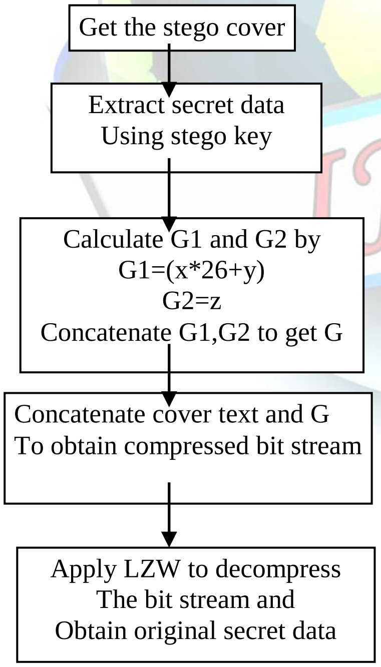
Concatenate G_1 and G_2 to obtain G.

Step 3:Concatenate the cover text and G to obtain the compressed secret data bit stream.

Step 4: Apply LZW to decompress the bit stream to get the original secret data.

The capacity is calculated by,

$$\text{Capacity} = \frac{\text{Number of bits in secret message}}{\text{Number of bits in cover message}}$$



The output is,

Enter secret message ‘hai’

The encoded bit is 104 97 105

The stego key is qaja

Cover text:’A high capacity text steganography using LZW compression’

Enter Stego key ‘qaja’

The original data is ‘hai’

Capacity is 0.53%

V. Conclusion

This work proposes a new technique for text steganography that customs LZW compression and LSB approach for hiding the secret data in the onward mail platform. There are numerous benefits of the proposed method. First, it performs better in terms of computational complexity, by this it saves the time required to process operation. Secondly, the usage of LZW directly on the secret data increases the embedding capacity. To further increase the capacity, LSB technique is used in the cover text or message of the to hide some part of the secret data bit stream. Security of the proposed method has been increased by employing stego keys. The proposed method can be applied to any language by reproducing the text directory and adapt the Latin Square to the corresponding language. Hence, it is not language explicit. The proposed method has supplementary advantages that it preserves the cover media while assigning and therefore it changes neither definition nor appearance of the cover text so the text is relevant and grammatically accurate and authorized. Compared to other techniques the experimental results of proposed data hiding scheme present the embedding capacity increased up to 13.43%. Therefore, this substantial performance improvement establishes the effectiveness of the projected algorithm.

REFERENCES

[1]. N. Johnson, Z. Duric, S. Jajodia, Information hiding, and watermarking-attacks countermeasures, in: Advances in Information Security, 2001.

[2]. F.K. Mohamed, A parallel block-based encryption schema for digital images using reversible cellular automata, Int. J. Eng. Sci. Technol. 17 (2014) 85–94.

[3]. C. Karri, U. Jena, Fast vector quantization using a Bat algorithm for image compression, Int. J. Eng. Sci. Technol. (2015).

[4]. W. Bender, D. Gruhl, N. Morimoto, A. Lu, Techniques for data hiding, IBM Syst. J. 35 (1984) 313–336.

[5]. Z.H. Wang, H.R. Yang, T.F. Cheng, C.C. Chang, A high-performance reversible data-hiding scheme for LZW codes, J. Syst. Software. 86 (2013) 2771–2778.

[6]. T.A. Welch, A technique for high-performance data compression, IEEE. 17 (1984) 6–17.

[7]. D. Salomon, Data Compression, Springer-Verlag, 2002.

[8]. E. Satir, H. Isik, A compression-based text steganography method, J. Syst. Software. 85 (2012) 2385–2394.

[9]. P. Wayner, Mimic functions, Cryptologia 16 (1992) 193–214.

[10]. K. Winstein, Lexical steganography through adaptive modulation of the word choice hash, secondary education at the Illinois Mathematics and Science Academy http://alumni.imsa.edu/keithw/tlex/lsteg.ps1999.

[11]. B. Murphy, C. Vogel, The syntax of concealment reliable methods for plain text information hiding, in: Proceedings of the SPIE International Conference on Security, Steganography and Watermarking of Multimedia Contents, vol. 6505, 2007.

[12]. H. Nakagawa, K. Sampei, T. Matsumoto, S. Kawaguchi, K. Makino, I. Murase, Text information hiding with preserved meaning—a case for Japanese documents, IPSJ Trans. 42 (2001) 2339–2350.

[13]. X.M. Sun, G. Luo, G. Huang, Component-based digital watermarking of Chinese texts, in: Proceedings of the 3rd International Conference on Information Security, 2004, pp. 76–81.

[14]. Z. Wang, C. Chang, C. Lin, M. Li, A reversible information hiding scheme using left-right and up-down Chinese character representation, J. Syst. Software. 82 (2009) 1362–1369.

[15]. Z.H. Wang, T.D. Kieu, C.C. Chang, M.C. Li, Emoticon-based text steganography in chat, in: Proceedings of Asia-Pacific Conference on Computational Intelligence and Industrial Applications, vol.2,2009, pp. 457–460.

[16]. R. Stutsman, M. Atallah, C. Grothoff, K. Grothoff, Lost in just the translation, in: Proceedings of the ACM Symposium on Applied Computing, 2006, pp. 23–27.

[17]. N. Samphaiboon, Steganography via running short text messages, Multimed Tool Appl. 52 (2011) 569–596.

[18]. A. Desoky, Listega: list-based steganography methodology, Int. J. Inf. Secur. 8 (2009) 247–261.

[19]. R. Kumar, S. Chand, S. Singh, An efficient text steganography scheme using Unicode Space Characters, Int. J. Forensic Comput. Sci. 10 (2015) 8–14.

[20]. R. Kumar, S. Chand, S. Singh, An Email based high capacity text steganography scheme using combinatorial compression, in: 5th IEEE International Conference CONFLUENCE 2014: The Next Generation Information Technology Summit, 25th-26th Sept., 2014, pp. 336–339.

[21]. R. Kumar, S. Chand, S. Singh, A high capacity Email based text steganography scheme using Huffman compression, in: International Conference on Signal Processing & Integrated Networks, 2016.

[22]. K. Tutuncu, A.A. Hassan, New approach in E-mail based text steganography, Int. J. Intell. Syst. Appl. Eng. 3 (2015) 54–56.

[23]. A.A. Mohamed, An improved algorithm for information hiding based on features of Arabic text: a Unicode approach, Egypt. Inf. J. 15 (2014) 79–87.