VOL. 14, NO. 1, JANUARY 2019 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

SPACE EFFICIENT IMAGE STEGANOGRAPHY: A NOVEL DATA HIDING SCHEME FOR TAMIL TEXT DOCUMENT

K. Manimozhi, S. Abiramasundari, V. Kalaichelvi, H. Manikandan and P. Meenakshi
Department of Computer Science Engineering, SASTRA Deemed University, Thanjavur, Tamil Nadu, India
E-Mail: kmanimozhi77@src.sastra.edu

ABSTRACT

Hiding secret information in a cover, which is of same type or different type is called Steganography. The main goal of the Steganography is that the sharing of secret information should not be shown to the intermediaries during communication. Several algorithms have been developed having different kinds of cover media, such as text, image, audio or video. We propose an approach to convert a secret message, which is available as Tamil text document into a gray-scale image that can be created of any size based on the size of the secret message by the sender, and that is transmitted over communication channel. At the hiding end, the characters in the tamil text document are encoded after encryption, then the image is scanned row by row, and in each non black pixel four characters are hidden. Only if the color of a pixel is not black, then the characters are hidden into it. At the extraction end, the characters are extracted after finding that the pixel is not in black color. This proposed method shows an agreeable experimental result with the cover image chosen. The efficiency of this proposed method is that the size of the stego image is lesser than the size of the secret message. Security can be increased by encrypting the secret text before embedding into the cover image.

Keywords: information security, image steganography, encryption, compression.

INTRODUCTION

Steganography is the form of security technique for ensuring the secret messages unread. The main objective of using steganography is to embed and hide the important information in a cover text/image. This can be done by various encoding techniques. However, the secret communication can be achieved through cryptography and steganography as well. Nowadays, communication through medium is more wide-ranging that necessitates hide secret information.

Steganography can be broadly categorized into linguistic steganography and technical steganography in which technical steganography is classified into digital images, video, audio and text steganography. Image steganography embeds the given information into a cover image. Even though many different image steganography algorithms have been developed, the breaking of secret messages still happens. Therefore, an urge of developing the effective algorithms is a great challenge for the developers. In connection to that, developers either derive a new algorithm or upgrade the existing one.

The secret messages can be made invisible by tattooing inside the cover image/text through steganography leaving without any trace of suspension [1-3]. Anyone suspect that the image could have any secret messages results in the failure of steganography technique used [4]. The tamil text document can also be used for hiding as cover text in text steganography [5].

The prisoner's rescue problem illustrates the concept of advanced representation of steganography [6]. Tremendous approach of steganography methods have been used in images [2-8], video frames [9-10] and audio [11-12]. Since text steganography is always being a great challenge for the cryptanalysis [13].

Context free grammars of rewriting rules are used as a base for structuring the text in syntactical steganography which are syntactically accurate [14]. NICETEXT algorithm [15-16] is one such algorithm that is used to identify the underlying information in a cover file. It applies the part-of-speech for generating the words for a sentence. In some cases, the result of NICETEXT would yield meaningless sentences.

In lexical steganography, the secret message is hidden by the relevant words of natural language. It chooses a dictionary word much similar to the original word without making the word suspicious and it is applied to the sentence.

Information hiding can also be done by picking the first letter of the alternate text. This approach is also extended to selected words [17]. Irrelevant text can also be used to hide the secret message in a HTML file [20]. In line shifting method, a line is rotated to some degree α. Whereas in word shifting method the information is stored by altering the gap between the words and moving words [18-21].

The other methods include feature coding method, which enforces the method of altering the complete formation of the text for steganography [22], [23]. The white spaces in between the characters of text is used for hiding the information that are mainly applied for Indian languages [24-26] using the technique of Open Spaces method, feature coding method and dynamic programming method. The various symmetry of reflection and shapes of the characters in the line of text are also considered for hiding the secret data bits [27-29].

PROPOSED WORK

We propose an image steganographic technique for Tamil text document. The Tamil text comprises of vowels, consonants, vowel consonants and some punctuators that are listed. The text is encoded using the look up table given. We take a secret Tamil text document and a 32-bit PNG cover image as input. The prerequisite to hide the secret text into the cover image is that the cover

image must have at least one fourth of the text length pixels.

At the hiding end, once the text and suitable image is chosen, the text can be hidden into the image. The text is segmented into number of blocks based on the block size. The block size is n x n, where n must be an even number and it may range from two to 16. When the text is to be hidden, the suitable block size is found based on the text length, it is ensured that the text length is a multiple of chosen block size n x n. If not, then the excessive characters can be grouped into block of nearest even size. Padding also can be used if necessary.

Once blocks are ready, we apply folding technique block by block. One block is taken at a time as a table that is horizontally folded based on the midway. Horizontally folded table is then folded vertically. After both the folding are applied, the size of the block is reduced to n/2 x n/2.

After folding, each cell in the table will have four characters. We encode all the characters in the table with the help look up tables such as Table-1, Table-2, Table-3 and Table-4.

Once the text is encoded, it can be hidden into the cover image using the Hiding algorithm.

At the extraction end, the stego image is received and all the non black pixels are scanned and the codes are extracted, decoded, then the folding techniques are applied reversely to get the hidden secret message. The proposed mechanism is shown in Figure-1.

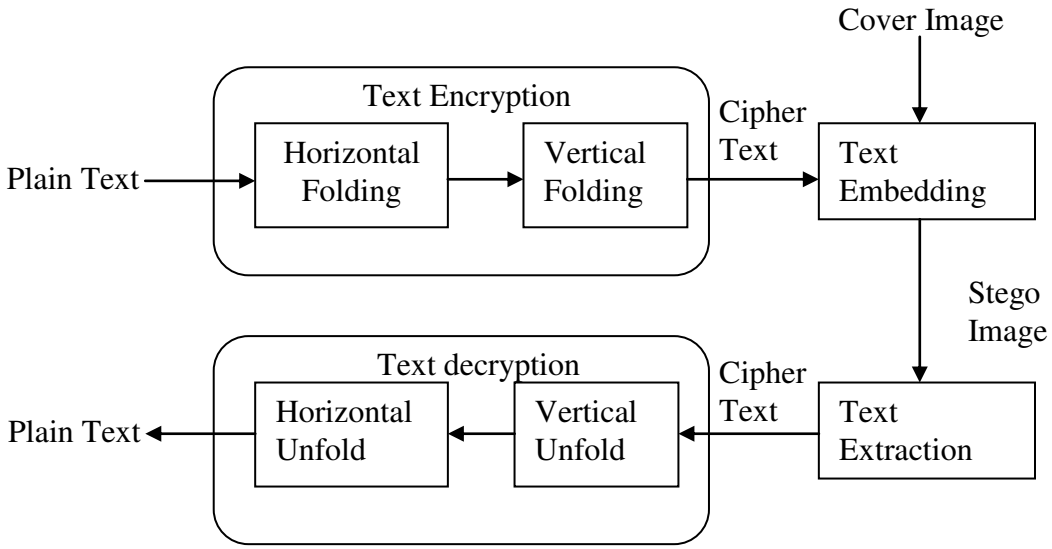

Figure-1. Proposed architecture.

Hiding algorithm

Input: Secret tamil text and a cover image (32-bit PNG image)

Steps:

1. Find the suitable block size 'n' such that it is suitable for folding.
2. If the text length is not a multiple of block size, then round it with the help of padding.
3. Take one block of characters from the input tamil document and store it in n x n table.
4. Apply horizontal folding so that the table is resized to n/2 x n and each cell contains two characters.
5. Apply vertical folding so that the table is resized to n/2 x n/2 and each cell contains four characters.
6. Encode the characters in the table by using the look_up table.
7. Take the cover image.
8. Apply raster scanning onto the cover image.
9. When a non black pixel is found, replace its byte values with next available cell's content.
10. Repeat the steps 8 and 9 until all the characters in the cells of the table are hidden.
11. Return the stego image.

Output: Stego image in which the secret tamil text is hidden

Extraction algorithm

The reverse process of the hiding algorithm is applied for extracting the secret tamil text from the stego image.

Table-1. Code for Tamil vowels.

1	2	3	4	5	6	7	8	9	10	11	12	0
அ	ஆ	இ	ஈ	உ	ஊ	எ	ஏ	ஐ	ஒ	ஓ	ஔ	ஃ

Table-2. Code for Tamil vowel consonants.

	அ	ஆ	இ	ஈ	உ	ஊ	எ	ஏ	ஐ	ஒ	ஓ	ஔ
க்	31	49	67	85	103	121	139	157	175	193	211	229
ங்	32	50	68	86	104	122	140	158	176	194	212	230
ச்	33	51	69	87	105	123	141	159	177	195	213	231
ஞ்	34	52	70	88	106	124	142	160	178	196	214	232
ட்	35	53	71	89	107	125	143	161	179	197	215	233
ண்	36	54	72	90	108	126	144	162	180	198	216	234
த்	37	55	73	91	109	127	145	163	181	199	217	235
ந்	38	56	74	92	110	128	146	164	182	200	218	236
ப்	39	57	75	93	111	129	147	165	183	201	219	237
ம்	40	58	76	94	112	130	148	166	184	202	220	238
ய்	41	59	77	95	113	131	149	167	185	203	221	239
ர்	42	60	78	96	114	132	150	168	186	204	222	240
ல்	43	61	79	97	115	133	151	169	187	205	223	241
வ்	44	62	80	98	116	134	152	170	188	206	224	242
ழ்	45	63	81	99	117	135	153	171	189	207	225	243
ள்	46	64	82	100	118	136	154	172	190	208	226	244
ற்	47	65	83	101	119	137	155	173	191	209	227	245
ன்	48	66	84	102	120	138	156	174	192	210	228	246

Table-3. Code for Tamil consonants.

13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
க்	ங்	ச்	ஞ்	ட்	ண்	த்	ந்	ப்	ம்	ய்	ர்	ல்	வ்	ழ்	ள்	ற்	ன்

Table-4. Code for punctuators.

247	248	249	250	251	252	253	254	255
Blank space	.	,	!	?	"	(	)	New line

EXPERIMENTAL RESULT

A sample Tamil text without any punctuators shown in Table-5 is taken and filled in a 16 x 16 block, Table-6 is the output of horizontal folding and Table-7 shows the output of vertical folding. At last the characters shown is Table-7 are encoded using the codes given in the Tables Table-1, Table-2, and Table-3. The encoded result is shown in Table-8. The encoded data is then hidden into the 32-bit PNG cover image given in Figure-1 and the stego image given in Figure-2 is created after hiding the encoded data.

Table-5. The secret Tamil text.

ந	ம	சி	வா	ய	வா	ழ்	க	நா	த	ன்	தா	ள்	வா	ழ்	க
இ	மை	ப்	பொ	ழு	து	ம்	எ	ன்	நெ	ஞ்	சி	ல்	நீ	ங்	கா
தா	ன்	தா	ள்	வா	ழ்	க	கோ	க	ழி	ஆ	ண்	ட	கு	ரு	ம
ணி	த	ன்	தா	ள்	வா	ழ்	க	ஆ	க	ம	ம்	ஆ	கி	நி	ன்
று	அ	ண்	ணி	ப்	பா	ன்	தா	ள்	வா	ழ்	க	ஏ	க	ன்	அ
நே	க	ன்	இ	றை	வ	ன்	அ	டி	வா	ழ்	க	வே	க	ம்	கெ
டு	த்	தா	ண்	ட	வே	ந்	த	ன்	அ	டி	வெ	ல்	க	பி	ற
ப்	ப	று	க்	கு	ம்	பி	ஞ்	ஞ	க	ன்	த	ன்	பெ	ய்	க
ழ	ல்	க	ள்	வெ	ல்	க	பு	ற	ந்	தா	ர்	க்	கு	ச்	சே
யோ	ன்	த	ன்	பூ	ங்	க	ழ	ல்	க	ள்	வெ	ல்	க	க	ர
ங்	கு	வி	வா	ர்	உ	ள்	ம	கி	ழு	ம்	கோ	ன்	க	ழ	ல்
க	ள்	வெ	ல்	க	சி	ர	ம்	கு	வி	வா	ர்	ஓ	ங்	கு	வி
க்	கு	ம்	சீ	ரோ	ன்	க	ழ	ல்	வெ	ல்	க	ஈ	ச	ன்	அ
டி	போ	ற்	றி	எ	ந்	தை	அ	டி	போ	ற்	றி	தே	ச	ன்	அ
டி	போ	ற்	றி	சி	வ	ன்	சே	வ	டி	போ	ற்	றி	நே	ய	த்
தே	நி	ன்	ற	நி	ம	ல	ன்	அ	டி	போ	ற்	றி	சி	வ	ம்

Table-6. Text after horizontal fold.

நதே	மநி	சின்	வாற	யநி	வாம	ழ்ல	கன்	நாஅ	தடி	ன்போ	தாற்	ள்றி	வாசி	ழ்வ	கம்
இடி	மைபோ	ப்ற்	பொறி	ழுசி	துவ	ம்ன்	எசே	ன்வ	நெடி	ஞ்போ	சிற்	ல்றி	நீநே	ங்ய	காத்
தாடி	ன்போ	தாற்	ள்றி	வாஎ	ழ்ந்	கதை	கோஅ	கடி	ழிபோ	ஆற்	ண்றி	டதே	குச	ருன்	மஅ
ணிக்	தகு	ன்ம்	தாசீ	ள்ரோ	வான்	ழ்க	கழ	ஆல்	கவெ	மல்	ம்க	ஆஈ	கிச	நின்	ன்அ
றுக	அள்	ண்வெ	ணில்	ப்க	பாசி	ன்ர	தாம்	ள்கு	வாவி	ழ்வா	கர்	ஏஓ	கங்	ன்கு	அவி
நேங்	ககு	ன்வி	இவா	றைர்	வஉ	ன்ள்	அம	டிகி	வாழு	ழ்ம்	ககோ	வேன்	கக	ம்ழ	கெல்
டுயோ	த்ன்	தாத	ண்ன்	டபூ	வேங்	ந்க	தழ	ன்ல்	அக	டிள்	வெவெ	ல்ல்	கக	பிக	றர
ப்ழ	பல்	றுக	க்ள்	குவெ	ம்ல்	பிக	ஞ்பு	ஞற	கந்	ன்தா	தர்	ன்க்	பெகு	ய்ச்	கசே

Table-7. Text after vertical fold.

நதேகம்	மநிழ்வ	சின்வாசி	வாறள்றி	யநிதாற்	வாமன்போ	ழ்லதாற்	கன்ள்றி
இடிகாத்	மைபோங்ய	ப்ற்நீநே	பொறில்றி	ழுசிசிற்	துவஞ்போ	ம்ன்சிற்	எசேல்றி
தாடிமஅ	ன்போருன்	தாற்குச	ள்றிடதே	வாஎண்றி	ழ்ந்ஆற்	கதைண்றி	கோஅடதே
ணிக்ன்அ	தகுநின்	ன்ம்கிச	தாசீஆஈ	ள்ரோம்க	வான்மல்	ழ்கம்க	கழஆஈ
றுகஅவி	அள்ன்கு	ண்வெகங்	ணில்ஏஓ	ப்ககர்	பாசிழ்வா	ன்ரகர்	தாம்ஏஓ
நேங்கெல்	ககும்ழ	ன்விகக	இவாவேன்	றைர்ககோ	வஉழ்ம்	ன்ள்ககோ	அமவேன்
டுயோறர	த்ன்பிக	தாதகக	ண்ன்ல்ல்	டபூவெவெ	வேங்டிள்	ந்கவெவெ	தழல்ல்
ப்ழகசே	பல்ய்ச்	றுகபெகு	க்ள்ன்க்	குவெதர்	ம்ல்ன்தா	பிகதர்	ஞ்புன்க்

Table-8. Encoded data.

38,163,31,22	**40,74,27,44**	**69,30,62,69**	**62,47,28,83**	**41,74,55,29**	**62,40,30,219**	**27,43,55,29**	**31,30,28,83**
3,71,49,19	184,219,14,41	21,29,92,164	201,83,25,83	117,69,69,29	109,44,16,219	22,30,69,29	7,159,25,83
55,71,40,1	30,219,114,30	55,29,103,33	30,83,35,163	62,7,18,83	27,20,2,29	31,181,18,83	211,1,35,163
72,13,30,1	37,103,74,30	30,22,67,33	55,87,2,4	29,222,22,31	62,30,40,25	27,31,22,31	31,45,2,4
119,31,1,80	1,28,30,103	18,152,31,14	72,25,8,11	21,31,31,24	57,69,27,62	30,42,31,24	55,22,8,11
164,14,139,25	31,103,22,45	30,80,31,31	3,62,170,30	191,24,31,211	44,5,27,22	30,28,31,211	1,40,170,30
107,221,47,42	19,30,75,31	55,37,31,31	18,30,25,25	35,131,152,152	170,14,71,28	20,31,152,152	37,45,25,25
21,45,31,159	39,25,23,15	119,31,147,103	13,28,30,31	103,152,37,24	22,25,30,55	75, 31,37,24	16,111,30,13

Figure-2. Cover image.

Figure-3. Stego image.

PERFORMANCE ANALYSIS

An algorithm that is newly proposed must be analyzed for its performance. Any algorithm's efficiency is analyzed by two important factors such as Running time and memory space. Since this algorithm is applied in the network communication, both time and space are significantly important. Time efficiency of the hiding procedure is *linear* under best case. The hiding procedure includes folding, encoding and embedding techniques. The folding technique uses only the secret text. When we take the text length as n, the running time of folding technique is $O(n)$. After folding, the characters are encoded using the look up table. The best case running time of the encoding is $O(n)$, as each search takes $O(1)$ with any efficient data structure. The best case running time of the Embedding procedure is also $O(n)$, so that the best case total running time of the hiding procedure is linear. In worst case, the encoding procedure takes $O(n^2)$, since the worst case running time of each search is $O(n)$, so that the worst case total running time of the hiding algorithm is $O(n^2)$.

Generally the Space Efficiency deals about the amount of space occupied by an algorithm for its completion. This algorithm is space efficient since it uses no additional memory. Since four characters are hidden into a non black pixel, the size of the stego image is smaller than the size of the text document. A cover image with more non black pixels can be chosen to hide a document with more characters. Thus the proposed algorithms are a space efficient image steganographic technique.

CONCLUSION AND FUTURE ENHANCEMENT

An approach to convert a secret message, which is available as Tamil text document into a gray-scale image that can be created of any size based on the size of the secret message by the sender, and that is transmitted over communication channel. At the hiding end, the characters in the Tamil text document are encrypted and encoded, then the image is scanned row by row, and in each non black pixel four characters are hidden. Only if the color of a pixel is not black, then the characters are hidden into it. At the extraction end, the characters are extracted after finding that the pixel is not in black color. This proposed method shows an agreeable experimental result with the cover image chosen. The efficiency of this proposed method is that the size of the stego image is lesser than the size of the secret message. Security can be increased by encrypting the secret text before embedding into the cover image. In future, the security of the secret information can be enhanced by adding an efficient encryption technique.

REFERENCES

[1] C. Cachin. 1998. An Information-Theoretic Model for Steganography. In: proceeding 2nd Information Hiding Workshop. 1525: 306-318.

[2] R Chandramouli, N. Memon. 2001. Analysis of LSB Based Image Steganography Techniques. IEEE. pp. 1019-1022.

[3] N.F. Johnson, S. Jajodia. 1998. Staganalysis: The Investigation of Hiding Information. IEEE. pp. 113-116.

[4] D. Artz. 2001. Digital Steganography: Hiding Data within Data. IEEE Internet Computing. pp. 75-80.

[5] K. Manimozhi. 2015. An Approach for Text Steganography: Generating Tamil Text Summary Using Tamil Phonetics. International Review on Computers and Software (I.RE.CO.S.). 10(2).

[6] G. Simmons. 1983. The prisoners problem and the subliminal channel. CRYPTO. pp. 51-67.

[7] J. Chen, T. S. Chen, M. W. Cheng. 2003. A New Data Hiding Scheme in Binary Image. In: Proc. Fifth Int. Symp. on Multimedia Software Engineering. Proceedings. pp. 88-93.

[8] M. Shobana, P. Gitanjali, M. Rajesh, R. Manikandan. 2013. A Novel Approach for Hiding Image Using Pixel Intensity. International Review on Computers and Software (IRECOS). 8(4).

[9] G. Doërr and J.L. Dugelay. 2003. A Guide Tour of Video Watermarking. Signal Processing: Image Communication. 18(4): 263-282.

[10] G. Doërr and J.L. Dugelay. 2004. Security Pitfalls of Frameby- Frame Approaches to Video Watermarking. IEEE Transactions on Signal Processing, Supplement on Secure Media. 52(10): 2955-2964.

[11] K. Gopalan. 2003. Audio steganography using bit modification. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP '03). 2: 421-424.

[12] Valarmathi Ramakrishnan, G. M. Kadhar Nawaz. 2014. An Enriched Audio Steganography for Secret Message Communication Using Novel Embedding Technique. International Review on Computers and Software (IRECOS). 9(10).

[13] J.T. Brassil, S. Low, N.F. Maxemchuk and L.O'Gorman. 1995. Electronic Marking and Identification Techniques to Discourage Document Copying. IEEE Journal on Selected Areas in Communications. 13(8): 1495-1504.

[14] P. Wayner. 1992. Mimic functions. Cryptologia XVI, pp. 193-214.

[15] M. T. Chapman. 1997. Hiding the hidden: A software system for concealing ciphertext as innocuous text. Master's thesis, University of Wisconsin-Milwaukee.

[16] Peng Meng, Liusheng Huang, Zhili Chen, Wei Yang, Dong Li. 2008. Linguistic Steganography Detection Based on Perplexity. International Conference on Multi Media and Information Technology. pp. 217-220.

[17] T. Moerland. 2003. Steganography and Steganalysis. www.liacs.nl/home/tmoerlan/privtech.pdf.

[18] S.H. Low, N.F. Maxemchuk, J.T. Brassil and L. O'Gorman. 1995. Document marking and identification using both line and word shifting. Proceedings of the Fourteenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '95), 2-6 April, 2: 853-860.

[19] A.M. Alattar and O.M. Alattar. 2004. Watermarking electronic text documents containing justified paragraphs and irregular line spacing. Proceedings of SPIE - Volume 5306, Security, Steganography, and Watermarking of Multimedia Contents VI. pp. 685-695.

[20] K. Bennett. 2004. Linguistic Steganography: Survey, Analysis, and Robustness Concerns for Hiding Information in Text. Purdue University, CERIAS Tech. Report 2004-13.

[21] Y. Kim, K. Moon, and I. Oh. 2003. A Text Watermarking Algorithm based on Word Classification and Inter-word Space Statistics. Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR'03). pp. 775-779.

[22] K. Rabah. 2004. Steganography-The Art of Hiding Data. Information Technology Journal. 3(3): 245-269.

[23] Shirali-Shahreza M.H.; Shirali-Shahreza M. 2006. A New Approach to Persian/Arabic Text Steganography. Computer and Information Science, 2006. ICIS-COMSAR 2006. 5th IEEE/ACIS International Conference on 10-12 July 2006. pp. 310-315.

[24] D. Huang and H. Yan. 2001. Interword Distance Changes Represented by Sine Waves for Watermarking Text Images. IEEE Transactions on

Circuits and Systems for Video Technology. 11(12): 1237-1245.

[25] S. Changder, N.C. Debnath. 2008. An Approach to Bengali Text Steganography. Proceedings of the International Conference on Software Engineering and Data Engineering (SEDE-08), ISBN: 978-1-880843-67-3, pp. 74-78, Los Angeles, California, USA.

[26] S. Changder, N.C. Debnath, D. Ghosh. 2010. LCS based Text Steganography through Indian Languages. Proceedings of 2010 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT 2010), ISBN: 978-1-4244-5539-3, 8: 53-58, Chengdu, China.

[27] Shraddha Dulera, Devesh Jinwala, Aroop Dasgupta. 2011. Experimenting with the Novel Approaches in Text Steganography. International Journal of Network Security & Its Applications (IJNSA). 3(6).

[28] S. Changder, S. Das, D. Ghosh. 2010. Text Steganography through Indian Languages using Feature Coding Method. Proceedings of the 2nd International Conference on Computer Technology and Development (ICCTD), 2010, 10.1109/ICCTD.2010.5645849, pp. 501-505, November, Cairo.

[29] Anandaprova Majumdera, Suvamoy Changderb. 2013. A Novel Approach for Text Steganography: Generating Text Summary using Reflection Symmetry. International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA).