A New Approach to Text Steganography Using Adjectives

Tatwadarshi P. Nagarhalli [1]

P. G. Student, Department of Computer Engineering, Shree L. R. Tiwari College of Engineering, Mumbai,

India[1]

ABSTRACT: A sentence contains at least one clause. That is, a sentence contains at least one subject and at least one predicate. And to enhance the understanding of the sentence or give more information about the subject and the predicate, adjectives are used in a sentence. This paper proposes a new method at text steganography where the adjectives, which can be so easily used in an English language, are used to hide information.

KEYWORDS: Information Hiding, Text Steganography, Mahabharata, Adjective, Adjective Text Steganography, Mahabharata Text Steganography.

I. INTRODUCTION

The Mahabharata is one of the greatest and the longest epics ever written. This epic is made of one hundred thousand verses, making it an epic eight times longer than the Greek epics Iliad and Odyssey put together.

Sage Veda Vyasa, the writer of the epic, says that the Mahabharata is the encyclopaedia of virtues and other tenets. In fact, the Mahabharata itself says that "whatever is spoken about virtue, wealth, pleasure, and salvation may be seen elsewhere; but whatever is not contained in this is not to be found anywhere" [1]. It also believed that it contains information that is useful even in today's context. One such important context pertains to steganography.

Steganography is the art and science of hiding secret data in another medium. The data which has to be hidden is called secret message and the medium in which this secret message is hidden is called the cover document [2]. And since the eves dropper is unaware of the presence of the secret message the question of decoding minimizes.

This paper proposes a technique of hiding of the secret message or information by using an adjective or a group of adjectives that are present or can be seamlessly integrated into a grammatically correct English sentence.

II. RELATED WORK

Over the many decades now text steganography has been a very interesting field of study. And this interest has led to many new techniques being put forward by many scholars. In this section we will look at some of these methods pertaining to text steganography.

Moreland [3] put forward a technique of hiding of words in Specific Characters of Words. Here the secret word was hidden in a collection words of a paragraph, character by character, like the first character of every paragraph. Or be it the Line Shifting method [4] text steganography where the lines were shifted vertically by some degree, which was taken as 0 and 1.

Or the Random Character and Word Sequencing method [5]. Or, the Word Shifting method [5] [7], where the words were shifted horizontally to hide the information. Another way of hiding the information is by placing punctuation at proper places, this approach is called Syntactic method [8].

The Semantic method techniques [8] hides the message by using synonyms of some particular words, mostly nouns, adjectives, verbs and adverbs. Feature coding technique [9] of the text steganography information was hidden in the changes in feature or structure that were made. Or, be it the open spaces method [11] of the text Steganography

Also many other researches have been carried out in the field of text steganography recently including applying its use in chat [13] and SMS [14]. Also, emoticons have been used to hide the information [17].

The large number of dots in the Arabic, Persian and Urdu have prompted them being used for text steganography. The methods [9] [10] [15] exploits these dots to store information.

As far as the Mahabharata is concerned, mattitude of research has been carried out on its linguistic properties, on its virtues, its warfare techniques but very less on the context of cryptology and steganography techniques that it uses or have been made use of in the book.

Taking a leaf out of the Mahabharata, this paper presents a novel approach of using the adjectives to store hidden message or information.

III.   PROPOSED TECHNIQUE

In the Mahabharata at one instance Dronacarya, the teacher, wants to communicate that Arjun, the opponent warrior, has arrived on the battlefield to his colleague Bhisma without others in the hearing range knowing about it. So, instead of calling the name of the person directly Dronacarya uses many superficially unrelated adjectives. On the face of it the uses of adjectives were useless. But to the receiver, in this case Bhisma, it contained the meaning that the person that has come in front of them is no ordinary warrior but in fact it is Arjuna the great mythological warrior [16].

Taking this idea forward the paper proposes a technique of using the adjectives to store secret message. Here an adjective or a group of adjectives contain a particular secret message which will only be decoded by the intended receiver. At both the sender and the intended receiver they will have a database or a table mapping a particular adjective or a group of adjectives to a unique secret message that need to be sent over the unsecured channel. For example, the database or the table will be something like Table 1.

Table 1: Adjectives assigned to each secret message

Adjective	Secret Message
Abandoned	English is a good language
Adventurous	English can be a great language
Astonishing	Hindi is a good language
Astonishing, Adventurous	Hindi can be a better language
Dependent	Hindi will be a great language
Disgusting	Hindi is a great language
Disgusting, Astonishing	USA is super power
Abandoned, Dependent	Indian is an emerging Economy

Astonishing, Adventurous, Disgusting	India will become a super power
Astonishing, Adventurous, Disgusting, Dependent	India is on its way of becoming a super power

In Table 1 a secret message has been assigned to single adjectives or combination of adjectives. And, if you want to convey a message that 'Hindi will be a great language', the corresponding adjective is 'Dependent'. And, you could write 'Children's are dependent on their parents' as a cover sentence containing the adjective. The original secret message is completely hidden from the eve's dropper who might who might catch hold of your message.

IV.   ENCODING TECHNIQUE

For encoding or converting the secret message to a document containing the secret message the following process needs to be followed. The secret message that needs to be sent is searched from the table and the corresponding mapped adjectives are identified. These adjectives are used in a grammatically correct sentence. This is the document containing the secret message.
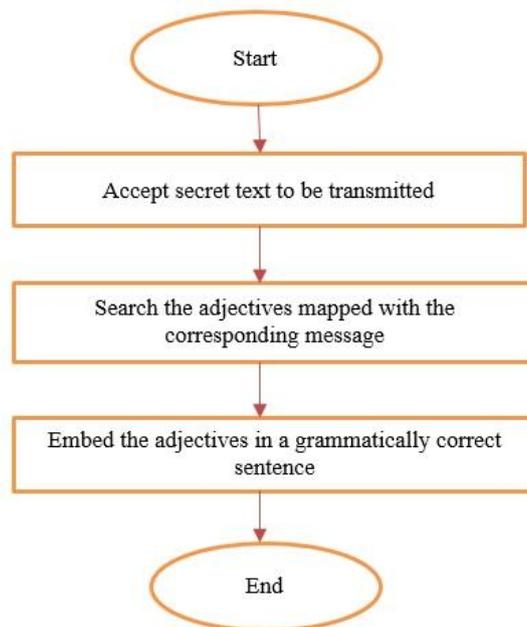

Fig 1: Flow chart for encoding the secret message

Fig 1 shows the encoding technique. First we need to accept the secret message that is to be transferred. The next step is to find the appropriate adjective corresponding to the secret message. And the final step is that we need to embed the adjective in a grammatically correct sentence. For example if the secret message is 'English is a good language', the corresponding mapped adjective is 'Abandoned'. The grammatically correct sentence will be 'His father abandoned him'.

V.   DECODING TECHNIQUE

For decoding or extracting the secret message the received message is scanned and the adjectives are tagged using natural language processing. The adjectives are extracted and mapped to the corresponding secret message associated with the concerned adjectives. This is the secret message sent by the sender in a secure manner by using the proposed technique.
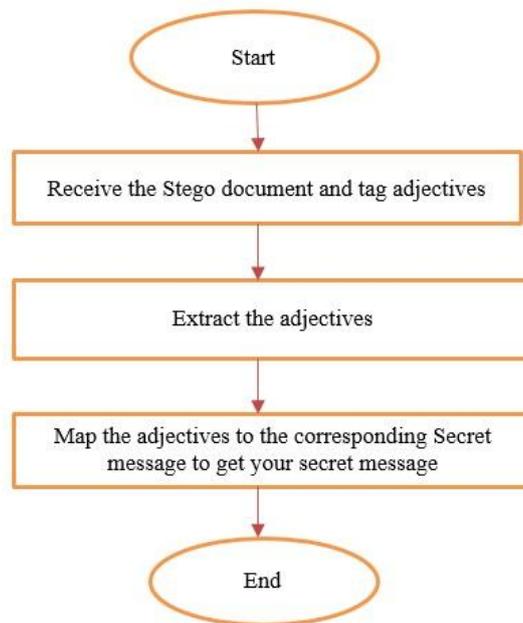

Fig 2: Flow chart for decoding the stego document

Fig 2 shows the decoding technique. Here the encoded message is received and the adjectives are tagged, these tagged adjectives are extracted and finally the adjectives are properly mapped for it corresponding secret message. For example here the received message is 'His father abandoned him', here the adjective is 'Abandoned' and the corresponding secret message is 'English is a good language'.

VI.   RESULTS

The simplicity of the proposed technique produces near accurate results. While preparing the grammatically correct sentence it should be made sure that extra adjectives are not inserted, as the system will extract all the adjectives that are present in the sentence for searching the appropriate secret message.

Table 2: Results with examples

Secret text to be hidden	Corresponding Adjectives	Cover text
English is a good language	Abandoned	His father abandoned him
Hindi is a great language	Disgusting	He has disgusting habit of drooling
Hindi is a good language	Astonishing	It was astonishing to see him alive

Hindi can be a better language	Astonishing, Adventurous	His astonishing and adventurous act paved the way for victory
Indian is an emerging Economy	Abandoned, Dependent	His father abandoned him when the son was most dependent on his father's advise

Table 2 shows the results with the secret text which was supposed to be hidden with its corresponding unique adjectives. The cover text was also prepared to show the completeness of the proposed technique.

VII.   ANALYSIS OF ADVANTAGES AND DISADVANTAGES

1. The technique is very easy to understand.
2. The implementation is also relatively simple, only the introduction of Natural Language processing for tagging adjective will have to be looked into.
3. Supposing that English language contains around Hundred thousand adjectives and scope of more, the secret message and the adjective pair that can be made is in the millions by just considering the possible combinations.
4. If permutation and combinations are taken into consideration, with unique order of the occurrence of the adjectives then the secret message and the adjective pairs that can be formed will be even more.
5. The database can be periodically updated without any hassles.
6. The ability of updating the database periodically, with ease, makes the system reliable and helps the technique be relevant for a long time.
7. The same technique of using the adjectives in the sentence can also be used and integrated to support any languages.
8. Instead of using synonyms which still might give away the secret message this technique hides the secret message more effectively.
9. One adjective can hide a complete sentence as a secret message rather than just 0 or 1.
10. The drawback of the system is that the generation of the database will be a somewhat tedious task.

VIII.   CONCLUSION

The paper proposes an easy method of hiding information making use of the adjectives, taking further the method or a way of hiding information used in the Mahabharata. These adjectives carry hidden information or secret message with it, unknown to a person to whom the message is not intended.

Also, the ease through which this technique of using the adjectives to covertly send the message, as used in the Mahabharata, has been integrated and carried further for technical use in this paper also proves that more research into the epic might throw some new techniques and methods in the field of modern steganography or even in many other different fields.

REFERENCES

[1] V. S. Sukthankar, "Adi Parva Adhyaya No. 56, Sholka No. 33", Bhandarkar Oriental Research Institute, The Mahabharata, Vol 1, Poona 1936, p. 244.
[2] M. Garg, A novel text steganography technique based on html documents, International Journal of advanced Science and Technology, vol.35, pp. 129-138, 2011.
[3] Bender, W. Gruhl, D. Morimoto, N. and Lu, A. Techniques for data hiding, IBM Systems Journal, vol. 35, Issues 3&4, pp. 313-336, 1996.
[4] S. H. Low, N. F. Maxemchuk, J. T. Brassil, and L. O'Gorman, Document Marking and Identification Using Both Line and Word Shifting, In 14th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '95), vol.2, pp. 853-860, 1995.

[5] A. M. Alattar, and O. M. Alattar, Watermarking electronic Text Documents Containing Justified Paragraphs and Irregular Line Spacing, In SPIE, Security, Steganography, and Watermarking of Multimedia Contents VI, vol. 5306, pp. 685-695, 2004.

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

[7] Y. Kim, K. Moon, and I. Oh, A Text Watermarking Algorithm based on Word Classification and Inter-word Space Statistics, In 7th International Conference on Document Analysis and Recognition (ICDAR03), pp. 775-779, 2003.

[8] M. Niimi, S. Minewaki, H. Noda, and E. Kawaguchi, A Framework of Text-based Steganography Using SD-Form Semantics Model, In Pacific Rim Workshop on Digital Steganography, Kyushu Institute of Technology, Kitakyushu, Japan, July 3-4, 2003.

[9] K. Rabah, Steganography-The Art of Hiding Data, Trans. Information Technology Journal, vol. 3, Issue 3, pp. 245-269, 2004.

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

[11] D. Huang, and H. Yan, Interword Distance Changes Represented by Sine Waves for Watermarking Text Images, Trans Circuits and Systems for Video Technology, IEEE, vol. 11, no. 12, pp. 1237-1245, 2001.

[12] M. Shirali-Shahreza, Text Steganography by Changing Words Spelling, In 10th International Conference on Advanced Communication Technology, ICACT 08, vol. 3, pp. 1912-1913, 2008.

[13] M. H. Shirali-Shahreza, M. Shirali-Shahreza, Text Steganography in Chat, In 3rd IEEE/IFIP International Conference in Central Asia, pp. 1-5, 2007.

[14] Z. Wang, et. Al, Emoticon-based Text Steganography in Chat, Second Asia-Pacific Conference on Computational Intelligence and Industrial Applications, pp. 457-460, 2009.

[15] J. A. Memon, K. Khowaja, and H. Kazi, "Evaluation of steganography for Urdu /Arabic text", Journal of Theoretical and Applied Information Technology, pp 232-237.

[16] V. S. Sukthankar, "Viraat Parva Adhyaya No. 683, Sholka No. 2", Bhandarkar Oriental Research Institute, The Mahabharata, Vol.5, Poona 1936, p. 180.

[17] T. P. Nagarhalli, "A new approach to SMS Text Steganography using Emoticons", IJCA Proceedings on National Conference on Role of Engineers in National Building, NCRENB: 1-3, June 2014.