# REVIEW PAPER ON STENOGRAPHY

## Pushpa Devi

*Ph.d student(computer applications)*

## ABSTRACT

*Steganography can be defined as a method of hiding data within a cover media so that other individuals fail to realize their existence. Image, audio and video are some popular media for steganography. But text is ideal for steganography due to its ubiquity and smaller size compared to these media. Particularly, the size rea lly does matter for mobile communication. However, text communication channels do not necessarily provide sufficient redundancy for covert communication. In this paper, a new approach for steganography in various financial statements is proposed. The main idea is that additional zeroes can be added before a number and also after the fractional part of a number without c hanging the value of the number*

## I. INTRODUCTION

Financial statements or finance al reports are formal records of the financial activities of a business, person, or other entity [1]. Income statements, statements of capital, balance sheets and cash flow statements are four common financial reports [2]. Financial statements are usually compiled on a quarterly and annual basis [3]. Financial statements are intended to be Understandable by readers who have " a reasonable knowledge of business and economic activities and accounting and who are willing to study the information diligently" [4] Steganography is the art and science of writing hidden messages in such a way that no one, apart from the sender and intended recipient, suspects the existence of the message, a form of security through obscurity [5]. Cryptography, the science of writing in secret codes, addresses all of the elements necessary for secure communication over an insecure channel, namely privacy, confidentiality, key exchange, authentication, and non repudiation. But, cryptography does not always provide safe communication. Consider an environment where the very use of encrypted messages causes suspicion. If a nefarious government or Internet service provider (ISP) is looking for encrypted messages, they can easily find them. Consider the following text file; what else is it likely to be if not encrypted? Steganography is the science of hiding information. Whereas the goal of cryptography is to make data unreadable by a third party, the goal of steganography is to hide the data from a third part

## II.PREVIOUS WORKS

Text steganography can be broadly classified into three types: format based, random and statistical generations, and linguistic method. Format based methods use and change the formatting of the cover text to hide data. They do not change any word or sentence, so it does not harm the 'value' of the cover text. Random and statistical

generation methods are used to generate cover text automatically according to the statistical properties of language. These methods use example grammars to produce cover text in a certain natural language. The linguistic method considers the linguistic properties of the text to modify it. The method uses linguistic structure of the message as a place to hide information. Following is the list of works has been done on hiding information or text steganography carried out.

1. Using markup languages : Some markup language feature can be used to hide information [7]. For instance, case insensitivity of HTML tags can be exploited. For example, the tag <BR> can be also used as <Br> and <br>. As a result one can do text steganography in HTML documents by changing the small or large case of letters in document tags. In some cases the positions of tags are also used. For example <B><U> </B></U> or like this <U><B> </U></B>. Information can be extracted by comparing the tags positions.

2. Using specific characters in text In this analytical, complicated and time consuming method, some specific characters from certain words are selected [8]. For example, the first words of each paragraph are sele cted in a manner that by placing the last characters of the selected words side by side forms the secret information.

3. Lines hifting and word shifting strategy In printed document the lines of the text are vertically shifted to some degrees [9], to hide secret data. Information can be also hidden in the text by shifting words horizontally and by changing distance between words [9], [11]. This method is acceptable for texts where the distance between words is varying. However, if the text is retyped or if character recognition programs (OCR) are used, the hidden information would get destroyed.

4. Exploiting punctuations igns Appropriate placement of some punctuation signs such as full stop (.) and comma (,) can hide information in a text file [8]. This method requires identifying proper places for putting punctuation signs

## III. PROPOSED METHOD

The idea behind our technique is very simple. We can pad arbitrary number of zeros before a number. We can also add arbitrary number of zeros after the fractional part of a number. Though padding any number of zeros is permitted, only one zero before a number is common. Similarly, in common practice two or one positions after the decimal point is significant. So we have four options: no padding, padding a zero before a number, padding one zero after the fractional part and padding two zeros after the fractional part of a number. With these alternatives we can represent two bits with a single case. All of the financial statements or reports contain huge number of numerical data. Padding some meaningless zero before or after the fractional part of a number can represent some information in any of the financial statements without drawing attention of people. In this paper we chose a sample balance sheet for hiding information. Another important thing to note is that though financial statements are generally formatted, that means unnecessary zeros before and after a number are omitted, unformatted financial statements are not very rare. More over, general users feel no curiosity about the unformatted financial statements, though they may grow a little bit annoyed. So unformatted financial

statements have a fair scope of being used as a cover media for steganography. Table-1 shows the summary of the hiding algorithm

Algorithm-1 and Algorithm-2 listed below summarize the encoding (or data hiding) and decoding (or extraction) procedure of the proposed method respectively. Figure 3 and Figure 4 displays the screen-shots of the encoder and the decoder of the Java application for the proposed steganography technique. The Java class *java.util.StringTokenizer* is used to tokenize the input financial statement and *java.util.regex* class is used to identify the numbers, which are the candidate tokens for data hiding, in the input financial statement.

**Algorithm-1: Hiding secret bits**

*Input*: Cover financial statement, secret bit stream

*Output*: Stego financial statement

*Step-1*: Make pairs of bits in the secret bit stream from right to the left. Pad a '0' before the stream if the number of bits is odd.

*Step-2*: Locate the first number in the financial statement. To search, move from left to right. If the current row is finished, then move on the next row until a number is found.

*Step-3*: For each pair of bit in the secret bit stream do

a) If the pair is 00 add an extra 0 after the fractional part of the number

b) If the pair is 01 add an extra 00 after the fractional part of the number

c) If the pair is 10 add an extra 0 before the integer part of the number

d) If the pair is 11 no processing of the number is required.

*Step-4*: Move on the next number as specified in step-2.

**Algorithm-2: Extracting secret bits**

*Input*: Stego financial statement

*Output*: extracted secret bit stream

*Step-1*: Locate the first number in the financial statement. To search, move from left to right. If the current row is finished, then move on the next row until a number is found.

*Step-2*: Look for unnecessary 0s in the number

a) If there is an unnecessary 0 after the fractional part of the number add 00 in the bit stream.

b)If there are two unnecessary 0 after the fractional part of the number add 01 in the bit stream.

c)If there is an unnecessary 0 before the integer part of the number add 10 in the bit stream.

d)If there is no unnecessary 0 before or after the number add 11 to the bit stream.

Step-3: Move on the next numbers as specified in step-1 and repeat step-2 and 3.
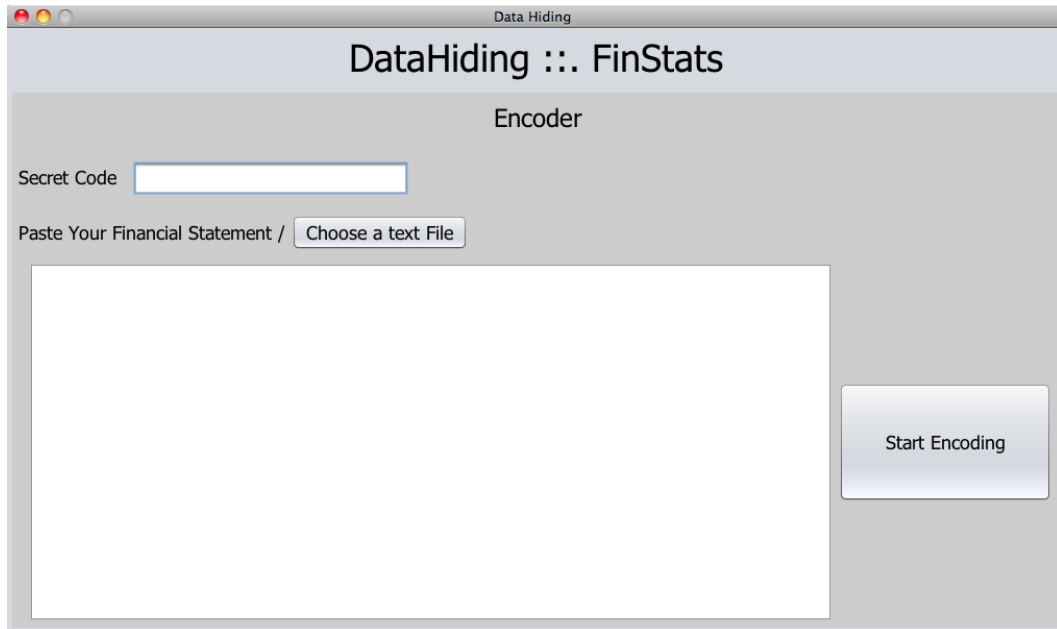
## IV. CONCLUSION

For text steganography various methods have been proposed. We represent a novel technique of hiding information in a financial statement. Financial statements are very much common in business and other organizations, which deal with finance. The capacity of this method is also in fair margin. We hide roughly 10 bytes of data in the sample balance sheet of Figure 1. Personal identification number (PIN), password, code, etc. are some commonly used secret information which are sent to the client via mobile SMS or by e-mail by the serving organization and generally have 8 to 12 characters. So, our proposed method can be a good choice for text steganography.
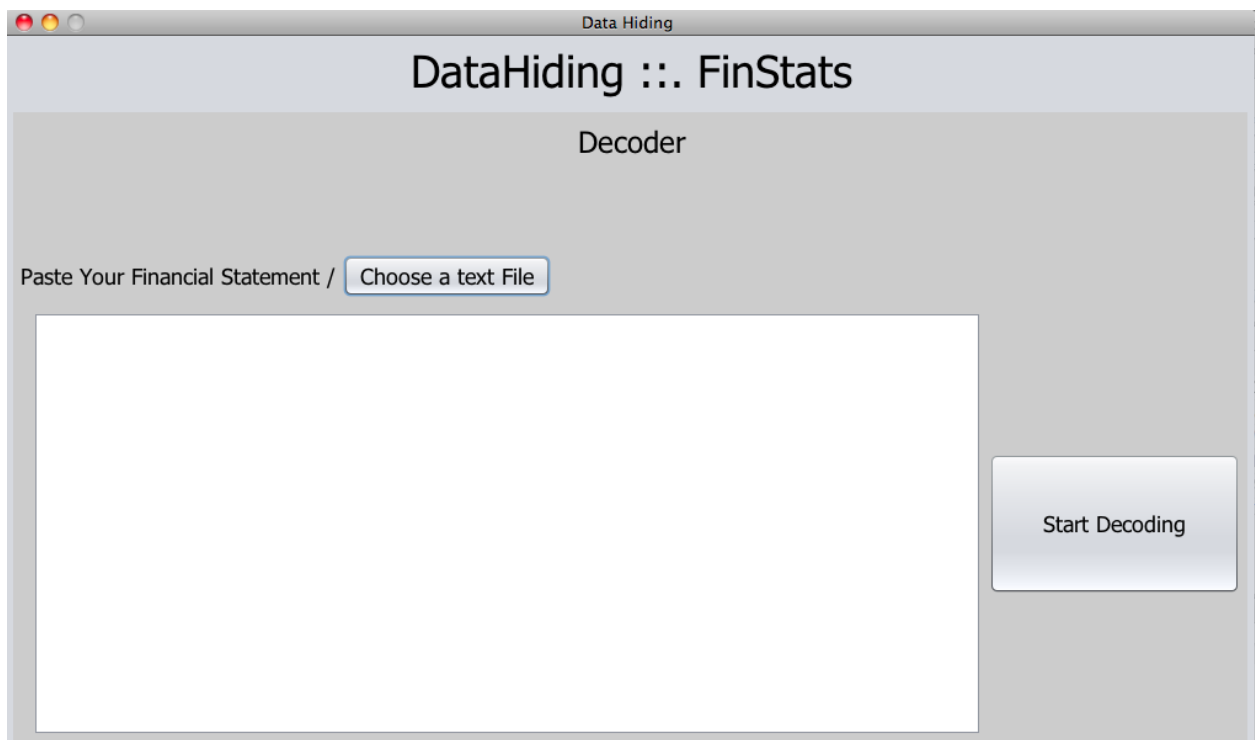
## REFERENCES

[1] "Financial statements", http://en.wikipedia.org

[2] "Common financial reports", http://allBusiness.com

[3] "Financial statement", http://www.investorwords.com/

[4] "The framework for the preparation and presentation of financial statements", http://www.iasplus.com/standard/framewk.htm

[5] "Steganography", http://en.wikipedia.org

[6] "Steganography: hiding data within data", http://sover.net

[7] K. Bennett, "Linguistic steganography: survey, analysis, and robustness concerns for hiding information in text", Purdue University, CERIAS Tech. Report-13 **(2004)**.

[8] T. Morkel, J.H.P. Eloff and M.S. Olivier, "An overview of image steganography", Proceedings of the Fifth Annual Information Security South Africa Conference (ISSA2005), Sandton, South Africa, Jun./Jul. **(2005)**, pp. 1-11.

[9] S. H. Low, N. F. Maxemchuk, J. T. Brassil, L. O'Gorman, "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), vol.2, Apr. **(1995)**, pp. 853 - 860.

[10] A.M. Alattar, O.M. Alattar, "Watermarking electronic text documents containing justified paragraphs and irregular line spacing", Proceedings of SPIE, Volume5306, Security, Steganography, and Watermarking of Multimedia Contents VI, Jun. **(2004)**, pp. 685-695.

[11] Y. Kim, K. Moon, and I. Oh, "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), **(2003)**, pp. 775–779.

[12] M. Niimi, S. Minewaki, H. Noda, E. Kawaguchi, "A framework of text-based steganography using sdform semantics model", Pacific Rim Workshop on Digital Steganography 2003, Kyushu Institute of Technology, Kitakyushu, Japan, Jul. 3-4, **(2003)**.

[13] K. Rabah, "Steganography-the art of hiding data", Information Technology Journal, vol .3, no. 3, **(2004)**, pp. 245-269.

[14] D. Huang, and H. Yan, "Inter-word distance changes represented by sine waves for watermarking text images", IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 12, Dec. **(2001)**, pp. 1237-1245.

[15] M. H. Shirali-Shahreza, and S. Shirali-Shahreza, "A new approach to persian/arabic text steganography", Proceedings of 5th IEEE/ACIS International Conference on Computer and Information Science and 1st IEEE/ACIS, Jun. **(2006)**.

[16] M.H. Shirali-Shahreza, and S. Shirali-Shahreza, "A robust page segmentation method for persian/arabic document", WSEAS Transactions on Computers, vol. 4, no. 11, Nov. **(2005)**, pp. 1692-1698.

[17] J.A. Memon, K. Khowaja, and H. Kazi, "Evaluation of steganography for urdu /arabic text", Journal of theoretical and applied information technology, vol. 4, no. 3, Mar. **(2008)**, pp. 232-237.

[18] M Aabed, S Awaida, AR Elshafei, A Gutub, "Arabic diacritics based steganography", IEEE International Conference on Signal Processing and Communications (ICSPC 2007), Dubai, UAE, Nov. **(2007)**, pp. 756-759,

[19] A Gutub, Y Elarian, S Awaida, A Alvi, "Arabic text steganography using multiple diacritics", WoSPA 2008 – 5th IEEE International Workshop on Signal Processing and its Applications, University of Sharjah, Sharjah, UAE, Mar. **(2008)**.

[20] B. Yang; X. Sun; L. Xiang; Z. Ruan, R. Wu, "Steganography in MS Excel document using text-rotation technique", Information Technology Journal, vol. 10, no. 4, **(2011)**, pp. 889-893.

[21] M. Khairullah, "A novel text steganography system using font color of the invisible characters in Microsoft Word documents", Proceedings of the 2nd International Conference on Computer and Electrical Engineering, Dubai, UAE, **(2009),** pp. 482-486.

[22] M. Khairullah, "A novel text steganography system in cricket match scorecard", International Journal of Computer Applications, New York, USA, vol. 21, no. 9, May **(2011)**, pp. 43-47.

[23] H. Kabetta, B.Y. Dwiandiyanta, and Suyoto, "Information hiding in CSS: a secure scheme text steganography using public key cryptosystem", Int. Journal on Cryptography and Information Security, vol. 1, **(2011)**, pp. 13-22.

[24] Y. Bassil, "A Generation-based Text Steganography Method using SQL Queries", International Journal of Computer Applications, vol. 57, no. 12, Nov. **(2012)**, pp. 27-31.

[25] R.K. Tiwari, "Microsoft Excel file: a steganographic carrier file", International Journal of Digital Crime and Forensics, vol. 3, no. 1, Jan.-Mar. **(2011)**, pp. 37-52.

**SNAPSHOTS**

.



The Encoder of the Developed Steganography Application



The Decoder of the Developed Steganography Application