Microsoft Excel File: A Steganographic Carrier File

Rajesh Kumar Tiwari, GLA Institute of Technology & Management, India
G. Sahoo, BIT Mesra, India

ABSTRACT

Data hiding is emerging as an important area of research with the advent of digital technologies. Watermarking, fingerprinting and steganography are important applications of data hiding techniques. Besides the classical use of data hiding in ownership protection, authentication and copy/access control, it serves as a general tool to convey side information. New files are a good choice for steganographic carriers. In this paper, the authors use Microsoft Excel as a cover file. A methodology has been developed that consists of three techniques for embedding the valuable data with three layers of security and high robustness. The resultant stego Excel file can be transmitted anywhere using any channel. Experimental results show the feasibility of the proposed methodology.

Keywords: Appending, Data Hiding, Encryption, Microsoft Excel, Steganography, Trash Space

1. INTRODUCTION

The digital information revaluation has brought significant changes in our life and society. Many advantage of digital information have generated new opportunities for innovation. Along with powerful software, new devices, such as digital camera, high quality printer and scanner, digital voice recorder, audio player etc. have reached consumer worldwide to create, manipulate and enjoy digital data. With the ease of editing of the digital data and also perfect reproduction in digital domain, the protection of ownership and the prevention of unauthorized tampering of digital data have raised serious concern. The common approach is to convert the digital data in different form so that the resultant data can be understood only by those who can get it back into its original form. This technique is known as encryption.

Data encryption has been proposed as a solution to save the privacy, security and to prevent illegal access of the data. The number theory based encryption techniques such as DES, IDEA, AES, RSA etc. have limitations when dealing with image or multimedia data. However, a major drawback of these techniques is that the existence of data is not hidden. Given enough time the unreadable encrypted data can be obtained in its original form.

As a part of solution, the concept of steganography has considerably given much attention in recent years. The word steganography is a Greek word which means to it as ‘writing in hiding’. The main purpose of steganography is to hide data in a cover media so that other will
not notice it (Venkatraman et al., 2004). The characteristics of cover media depends on the amount of data that can be hidden, the perceptibility of the message and its robustness (Amin et al., 2003; Marvel & Retter 1996; Swanson et al., 1996; Venkatraman et al., 2004).

In digital steganography, image areas that contain complex texture are candidate locations for carrying data since the human visual system is insensitive to high frequency or low changes in their resolution.

Fields like publishing and broadcasting fields also require an alternative solution for hiding information. Unauthorized copying becomes a hot issue in areas like music, film, print and software. To overcome such type of problems some invisible information can be embedded in the digital media in such a way that no one can easily extract it (Artz, 2001; Johnson & Jajodia, 1998; Anderson & Petitcoals, 1998; Craver, 1998; Sahoo & Tiwari, 2010a, 2010b).

2. LITERATURE REVIEW

Encoding secret message in text can be a challenging task. This is because text files have small amount of redundant data that is to be replaced with a secret message. Another drawback is the easy alteration of characters which can be done by unwanted parties. There are some methods by which we can accomplish text based steganography. These methods are:

- Open Space Method
- Syntactic Method
- Semantic Method

There are numerous ways to employ the open space in text files to encode the information. This method works because a casual reader may avoid one extra space at the end of line and an extra space between two words does not prompt abnormality. The extra space can be used to conceal one bit either 0 or 1. However, inter sentence space method encodes a “0” by adding a single space after a period in English prose. Adding two spaces would encode a “1”.

The left shift encoding involves actually shifting each line of text vertically up or down by as little as three centimeters. The End-of-line space method exploits white space at the end of each line.

Syntactic methods exploit the use of punctuation and structure of text to hide data without scientifically altering the meaning of the message. For example the two phrases “bread, butter, and milk” and “bread, butter and milk” are grammatically correct but differ in the use of comma. One can use this structure alternatively in a text message to represent either a “1” or “0”. Semantic methods assign two synonyms a primary or a secondary value. These values are then translated into binary “1” or “0”. For example, the word “big” is assumed primary and the “large” is assumed secondary. Therefore, decoding a message would translate the use of primary to be “1” and secondary to a “0”. The main hurdle is that replacement of the synonyms may change the meaning of the structure of the sentence.

Donovan Artz (2001) has designed a methodology for hiding the secret data by changing the order of characters or words. Ordering data that does not have an ordering constraint is often an effective method of steganography. Each permutation of a set of characters or words can be mapped to a positive integer. This mapping can then be used to encode hidden data by altering the order of characters or words.

Mathkour et al. (2008) have proposed a more robust steganographic technique that takes advantages of the strengths and avoid the limitations. They have given a detail study of the strength and weakness of the proposed techniques.

Liu and Tsai (2007) have proposed a new low embedding steganographic method for data hiding in Microsoft Word documents by a change tracking technique. The data embedding is disguised such that the stego document appears to be the product of a collaborative writing effort.

Shi and Zou (2005) have given a novel formatted text document data hiding algorithm, called inter-word space modulation (ISM)
Attack Graph Analysis for Network Anti-Forensics
www.igi-global.com/article/attack-graph-analysis-for-network-anti-forensics/110395?camid=4v1a