International Journal of Software Innovation
Volume 7 • Issue 3 • July-September 2019

Text Steganography Approaches Using Similarity of English Font Styles

Sahar A. El Rahman, Electrical Engineering Department, Faculty of Engineering-Shoubra, Benha University, Cairo, Egypt and Computer Science Department, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia

ABSTRACT

Data protection has become a more critical issue and the necessity to secure a transmission channel is become more serious. Therefore, steganography, the art of hidden data into a digital media in a way that embed a secret message in the cover document without permitting anyone to suspect the data existence except the intended recipient, has become a relevant topic of research. The actual challenge in steganography is how it could obtain high robustness and capacity without damaging the cover document imperceptibility. This article presents two steganography approaches that based on the Similarity of English Font Styles (SEFS). This process has the main document font style replaced by a similar font style to embed the secret message after encoding it. This is done by using 1) the upper-case letters and punctuation marks of the carrier document or 2) the white space between words, start and end letters of each word that has more than 2 letters in the carrier document. These approaches are tested by being applied to various document formats with various font styles. From the findings, the secret message was vague to an antagonist and the stego-document size was increased and the capacity is very high. Also, the approaches are implemented using C# to develop a tool that hides a critical data in text document and the same findings were achieved.

KEYWORDS

Cryptography, Data Hiding, Steganography Techniques, Text Steganography

1. INTRODUCTION

Steganography comes from the Greek words (στεγανό-ς, γραφ-ειν) meaning, "covered writing". In the past, human used a hidden ink or concealed tattoos to transfer stegnographic messages. This day, data processor and internet technology supply simple to utilize different channels of communications for steganography (Agarwa, 2013; Provos & Honeyman, 2003).

Steganography is the science and art of concealing communications stenographic process, consequently hides the concealed contents in unnoticeable covering media, thus it will not induce the suspiciousness of eavesdroppers. Cryptography, in contra to, wherever the enemies are permitted to intercept, detect and update the contents without being capable to break a specific security promise ensured by a crypto-system. Steganography objectives are to conceal message into other safe media in a manner which doesn't let the enemies at all to notice that there are other messages (Dunbar, 2002; Changder et al., 2010a; Ammar et al., 2010; El Rahman, 2016; El Rahman et al., 2016). The steganography feature over just cryptography is the hidden content does not draw alertness to recipients, enemy, or messengers (Samphaiboon & Dailey, 2008; Ankit, 2007). A comparison between steganography and cryptography is presented in Table 1 (El Rahman, 2015).

DOI: 10.4018/IJSI.2019070102

Table 1. A comparison between steganography and cryptography (El Rahman, 2015)

	Steganography	Cryptography
Definitions	Blotting out the existence of the message	Blotting out the concept of the message
Carriers	Any digital media	Commonly text-based
Keys	Option	Needful
Target	Secure communications	Information protection policy
Visible	Seldom	Forever

Steganography used in many applications to contain secrete communication, whereas cryptographic encryption algorithms aren't obtainable and also, powerful cryptography is impractical. In several instances, for instance, in martial application, even the information which two parties communicate could be of great interest. Besides, in the healthcare, and particularly the systems of medical imaging, might extremely well-being from data hiding algorithms (Changder et al., 2009; Changder et al., 2010a).

Steganography techniques are in common (see Figure 1) established on substituting noisy components of a digital medium by hidden messages. The system Security must not be established on an embedding technique, but on concealing the key (Changder et al., 2010b).

The "Embedded" data is the data to be concealed in the cover media. The data include the cover media and the "embedded" data together is known as the "Stego" data. Reasonably, the procedures of embedding the embedded, or the hidden data into the cover media, is often known as embedding. The expression "cover" is utilized to characterize the innocent message, original, data, video, audio, and so on, as shown in Figure 2. The procedure could be symbolized by Equation (1) (Channalli & Jadhav, 2009):

$$Stego\text{-}Medium = Embedded\ Message + Stego\text{-}Key + Cover\ Medium \quad (1)$$

2. STEGANOGRAPHY TYPES

2.1. Fragile

It involves hiding data in a file, that is ruined if the file has changed. This type is inappropriate for registration the copyright possessor of the file because it can be very easily taken away, however, it is beneficial in cases wherever it is serious to show that the file has not been meddled with, such as applying a file as proof in a court of jurisprudence, due to any tampering would take in the watermarking. Fragile algorithms tend to be easier to proceed than robust types (Cummins et al., 2004; Provos & Honeyman, 2003).

2.2. Robust

Its objective is embedding data into a file, that can't be put down completely. Though no scar is truly indestructible. If the sum of alterations desired to transfer the mark will render the file worthless, then, the system can be considered robust. Consequently, the mark should be veiled in a section of the file where its removal will be simply realized (Protection against detection), as depicted in Figure 3 (Cummins et al., 2004; Provos & Honeyman, 2003).