An overview and computer forensic challenges in image steganography

ABDULLAHI YARI, Imrana and ZARGARI, Shahrzad <http://orcid.org/0000-0001-6511-7646>

Published version

ABDULLAHI YARI, Imrana and ZARGARI, Shahrzad (2017). An overview and computer forensic challenges in image steganography. In: Proceedings, 2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). IEEE, 360-364. (In Press)

An Overview and Computer Forensic Challenges in Image Steganography

Imrana Abdullahi Yari
Department of Information System Security
Sheffield Hallam University
Sheffield, United Kingdom
imranayari@rocketmail.com

Shahrzad Zargari
Department of Information System Security
Sheffield Hallam University
Sheffield, United Kingdom
S.Zargari@shu.ac.uk

***Abstract*—The development of powerful imaging tools, editing images for changing their data content is becoming a mark to undertake. Tempering image contents by adding, removing, or copying/moving without leaving a trace or unable to be discovered by the investigation is an issue in the computer forensic world. The protection of information shared on the Internet like images and any other confidential information is very significant. Nowadays, forensic image investigation tools and techniques objective is to reveal the tempering strategies and restore the firm belief in the reliability of digital media. This paper investigates the challenges of detecting steganography in computer forensics. Open source tools were used to analyze these challenges. The experimental investigation focuses on using steganography applications that use same algorithms to hide information exclusively within an image. The research finding denotes that, if a certain steganography tool A is used to hide some information within a picture, and then tool B which uses the same procedure would not be able to recover the embedded image.**

***Keywords*-Image Steganography, LSB Steganography, LSB Algorithm of Spatial Domain, steganalysis**

I. INTRODUCTION

The world of computing era has led to digital media advancement and increase in widespread of computing options that are financially savvy or offer an efficient use. Those requiring power and rationale intending to resolve the issue in computer forensic examination tools and strategy that will enhance a robust computer forensic environment are of prime concern. Steganography is a technique used to hide information in a plain sight; it is like camouflage that could be invisible to the intruder or unintended recipient [1]. This study attempts to respond to the possibilities in computer forensic investigation by using tool X to decode the hidden information encoded by other tools which have the same features and follows same techniques. It is expected that tool X should be able to reveal the hidden information since both use same algorithms. It will be expressed in a practical context and the result is expected to open a room for further studies. If the results deviate from expectation, then it is a challenge in computer forensic investigation that it might possibly propose an alternative technique of decoding steganography.

The following section explores further understanding about different studies on the era of steganography and its detection techniques. Section 2 contains a brief discussion of steganography applications and techniques as surveyed from literature. The techniques employed by steganalysis will also be explored to build a foundation For any proposed solution. Section 3 involves the experimental work of decoding information encoded by steganography tools that uses the same procedure, followed by the findings and discussion leading to conclusion.

II. RELATED WORK

A. *Digital Era of Steganography*

The period of digital steganography plays a significant role in the realm of the digital world with the use of signal data processing programming and data theories [2]. The expanding technological innovation patterns of steganography used as a part of the different field like in networking, military, health, interactive media and so forth [3]. Moreover, the advancement of steganography is increasingly turning out to be where individuals are not just intrigued on concealing messages. Additionally, they are also willing to acquire the hidden data without twisting or removing the actual message in interactive media [4]. It was examined in the University of Michigan with around three million pictures from the cloud trying to find a trace to stenographic data, but they could not find a bit of any covert message; although evidences to the failed result was stated [5].

Steganography, watermarking and cryptology are interconnected as they are intended for secret communication. Watermarking is a sort of marker clandestinely inserted in a digital data as an image used for identification proof of ownership of such data [6]. Besides, it is opposed that steganography on its own does not provide integrity in privacy or encryption, but also suggest that combination of these functions can yield a stronger scramble information [7]. These findings show an evidence of a challenge in using a stenographic framework to protect information without integrating other functions like cryptography. The significance of cryptography is to cover up a message, rendering it incomprehensible without privacy and intension to be hidden as art of encryption that performs the transformation from plain text to cipher text [8]. However, steganography and watermarking shared the same feature of information hiding

technique. Intuitively, cryptology and steganography are families in such a way that cryptography scrambles message contents so it cannot be comprehended while steganography shrouds the existence of the message with the goal that it cannot be revealed [9].

B. The Steganography Detection Techniques

Steganalysis is an attack that aims to break steganography techniques in a fight that never ends. It is created to mirror cryptanalysis and also use stenographers in testing their algorithm quality instructions step by step to abstain from detection [10]. Steganalysis is developed with the use of several image processing strategies such as code translation. The attack is successful by observing the existence of hidden information in a file which is apparently different in a watermarking attack that is just to remove the watermark [11]. However, the recent development in steganography requires a strong technique to identify the hidden contents which have least false alert rate [4].

Additionally, the easiest and simplest way to detect or suspect the existence of steganography is by using the natural eye [12]. Experts in steganalysis observe the presence of steganography when each bit of pixels is altered [13]. The EncaseApps C-TAK is built with a dataset that helps in computer forensic investigation with an impact on the analysis in finding the accurate information in the examination of cyber threats and steganography. The accuracy part involves detecting even the specific type of steganography tool used for encoding [14]. This kind of tool is developed to investigate the known bad hash-sets that are integrated into datasets not with the outliers. A recent study suggests a new technique that can be incorporated into Encase forensic tool to detect and find the hidden information in a power point by using Encase Transcript [15].

III. Steganography Techniques and Algorithms

The process involves embedding the cover file from the sender and the convenient approach is applied at the expected beneficiary end to reveal the hidden message. Figure 1 below shows the techniques applied in steganography [16]. The strength of steganography depends on the following factors: the power of the secret data to remain hidden with the strong algorithm used, enough space to allocate the hidden data, the algorithm should robustly deliver the message safely from one end to the other without any data loss during compression and being resistant to attack during data transmission. Furthermore, the protection of the algorithm and passphrase used should be kept secret so that even if the attacker detects the presence of steganography, she/he cannot reveal the hidden data since the algorithm used is not exposed [5].

Steganography algorithms can be categorized based on the cover file used (image, text, audio, video, or protocol), the file format type (JPEG, BMP, or GIF) or method of compression used (e.g. JPEG: lossy or lossless), the domain type (transformed domain e.g. DCT method or spatial e.g. LSB Method), method used in embedding (Spread Spectrum, masking, statistical, or distortion) and so forth.

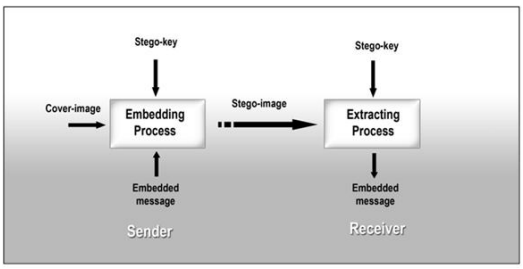

Figure 1. Steganography block diagram

A. Image Steganography

The hidden information is inserted as noise, which made it almost difficult to visualize by the naked human eyes. Images have a high level of redundancy and tolerance of twisting [17]. In deciding which algorithm method of steganography to implement, a type of compression technique is assumed to take a significant role. A JPEG image file format uses a lossy compression method that results in small image sizes with the chance that the concealed message might mostly be lost because much image data content will be deleted. While an example of lossless image compression technique is GIF. The lossless method does not compress the picture to its small size as lossy, but there is a high probability that digital image contents will not be lost. Steganography uses a redundant data within a picture content to store its secret data while the aim of image compression has the opposing purpose to steganography, which is to reduce the redundancy space in a picture so as to represent it in a lowest possible bit [18].

B. The LSB Algorithm of Spatial Domain

Least significant bits substitution (LSB) is the least type of algorithm in which LSBs of the covert information is altered and differs from a transformed domain with its high capability to allocate high space limit. Besides, steganalysis easily detects the concealed information that uses LSB. In transformed domain approach, the transformed variable stores the hidden data after initial transformation to a new domain and the original image is obtained from the edited image as a converse change to its original space. Discrete cosine transformation (DCT) is one of the widely used in this category [19]. LSB is done by changing the bits in the binary format of an image file in steganography [20]. Steganography is difficult to detect because some steganography tools that use LSB substitution for encoding consider changing the least bit whereas others randomize all the original bits in the cover file that is altered [21]. LSB involves changing the

bit of the image to store the secret information. Changing intensity is negligible but appears to be unchanged to the human eye. The hidden data will have no protection once it is discovered and the larger the image, the more prone to attack, due to their unusual size on transmission. Every image has a pixel which is responsible for a given colour of the picture. The pixel is represented in three primary forms of colour intensities R(red)G(green)B(blue). Depending on the intensity; LSB algorithm converts the data to binary with the last bit of the pixel [22].

IV. Experimental Work

This section attempts to examine practically with the use of computer forensic tools to find the challenges of computer forensic investigation in image steganography. There are many steganography tools, but only the ones that reflect our requirement for analysis were considered. In order to make a decision on which open source tools to be selected to run some experiments, further investigation was carried out and the results are demonstrated in table 1. [23] [24] [25] [26] [27] [28] [29].

Table I
ANALYSIS OF STEGANOGRAPHY TOOLS

Steganography tool	LSB method	Method of Encryption	Platform	Password support
JPHIDE (jphs)	√	blowfish	multi	√
SilentEye	√	AES	cross	√
S-tool	√	**DES**	**Windows**	√
OpenPuff		Joined Multi		√
OpenStego	√	**DES**	**cross**	√
QuickStego		none	Windows	

The S-tool and OpenStego steganography tools fit the requirement for this research as shown in the table above. However, there are still some little differences in terms of their support. The only surveyed key different feature between S-tool and OpenStego is that, S-tool supports more method of encryptions than OpenStego that supports only DES which have no impact on steganography process. For this experiment, two files were used for secret and cover file. The secret.txt file is containing a message that is intended to be secret, and the cover image file is cover.bmp. The file properties and the screen shots of all the process taken are shown in the below figure:

Figure 2 above illustrates the encoding method with a secret.txt and cover.bmp using S-tool. The text file with 81 bytes in size is embedded into a cover.bmp file using drag and drop to S-tool and passphrase. The passphrase used here is 2-0 and DES as the encryption method. The two images shown at the right represent the stego-file and the actual cover file. No difference in both files before and after steganography. The stego-image is saved as hidden.bmp. The StegExpose steganalysis tool is used to distinguish and

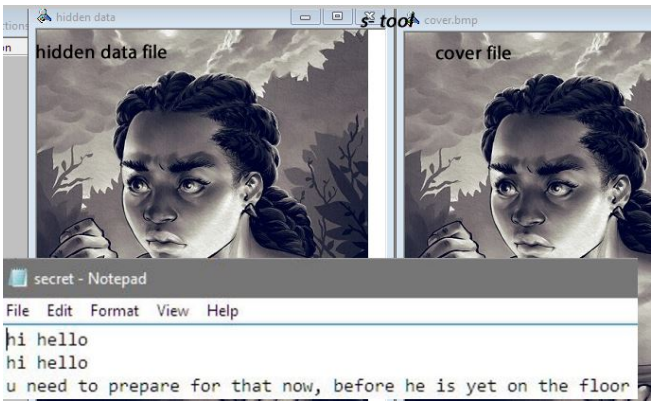

Figure 2. Encoding method: hiding process using S-tool

identify the presence of steganography file in a directory containing the stego-files as shown in figure 3 below, and the tool detected the hidden.bmp file as suspicious.

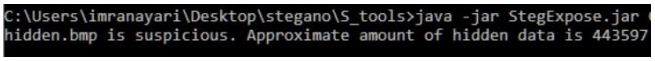

Figure 3. Detecting method: detecting process using StegExpose

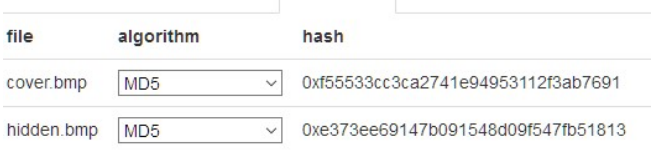

Figure 4. Online hexeditor: Comparing the files hashes

An online hex editor is used to compare and view the files. The figure 4 shows a difference between the files by comparing their hashes.

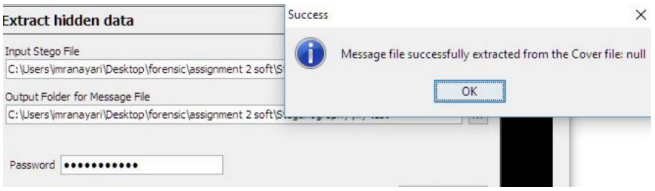

Figure 5. Decoding method: decoding method at the OpenStego

Figure 5 above shows an attempt to extract the hidden data using OpenStego tool that was stored by S-tool. The saved file is browsed and the same password was used to reveal the original data. The message shown at the pop-up indicates that the data has been extracted successfully. However, a notice message as shown in figure 6 from the java terminal indicates that there is an issue with the extraction. The message read, *embedded data is corrupt or invalid password has been provided or no algorithm found which can handle*

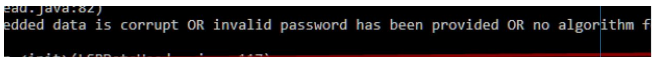

Figure 6. Generated error message

the given stego file. Due to this error, of course, no file is extracted

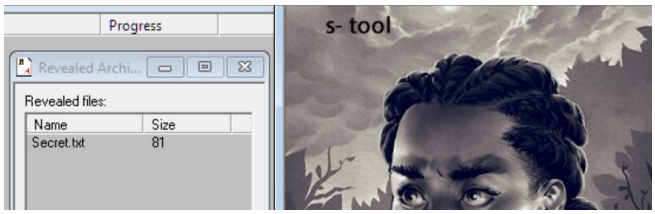

Figure 7. Decoding method: decoding method at the S-tool

Figure 7 illustrates the same stego-file that OpenStego failed to extract the hidden data was now extracted using a tool utilized in the encoding method. The password and encryption method used are all the same. The revealed files shown above have the same size as before encoded.

V. Analysis of Experiment Outcomes

It is expected that since both of the tools use the same techniques and the same password was used for encoding, the other tool should be able to decode and reveal the concealed information.

The outcomes of this experiment shows that it is not possible for a tool A in steganography to extract the data encoded by tool B even though they share the same techniques and features. The three possible errors for this are, the file might be corrupted on the way, which is not feasible since the S-tool was able to decode the message or perhaps the OpenStego sees the file as corrupt. Secondly, the probability that the password used is invalid, cannot be possible, since it was the same password used for encoding method. The possible error is an algorithm used that might cause an issue. However, both of these tools use LSB substitution method. The point here is that these tools might use a different way of choosing Least Significant Bit in their substitution method, possibly a randomization or last two digits or last digit only. This is the same idea expressed in Kessa, 2015 [21] study as reviewed in the literature section. The effects of corruption and unknown method of steganography are very difficult to recognise.

Thus, present a challenge in computer forensic investigation to find tools and techniques to break the concealed information in steganography when the need arise. It is not just to detect the presence of steganography but significantly need may arise to reveal hidden data information. This kind of generic tool development in steganography detection and classification is still developing.

VI. Conclusion

As steganography turns out to be used more broadly in digital world, there are many issues that should be known in computer forensic examination. There are wide assortments of various tools and techniques with their own focal points and weaknesses. Steady change should be made and more up to date adaptations. Initially, the overview of the digital era steganography gives an insight guidelines and understanding of steganography to the computer forensic expert in the field. The use of tools to observe the changes in bits of data may also trigger suspicions as surveyed. The procedures and algorithms used by steganography are analysed to serve as a basis for understanding how steganography works. Image file type is the most widely used medium of digital media, and this research is limited to LSB method since it is mostly used in digital image steganography and has little effect in altering the actual colour. This makes it hard to be detected by normal visualization. Moreover, the practical result clears a computer forensic expert thinking to use a steganography tool A for tool B to extract a hidden data, which is not possible as limited to this experiment, even though they share same properties and procedures.

Finally, this research shows a clue on how it is important for computer forensic examiners to know the type of steganography tool installed, hidden or deleted in the victim's computer. Finding evidence that the suspect uses a certain steganography tool triggers a dubious impression given that the victim uses steganography in the first place, thus evident a gap for subsequent investigation on finding the hidden files on the computer. Moreover, as shown in the experiment outcomes, knowing the type of steganography tool used is required to decode the hidden information, even though the tool utilized for the investigation have same features and follow same techniques with the victim's tool.

ACKNOWLEDGMENT

A considerable effort and time were applied in ensuring this research aim is achieved as stated. Moreover, I am expressing my sincere appreciation to Dr. Shahrzad Zargari for her support, guidance and patience in completing of this study. Her advice, assistance, and feedbacks are the sources of encouragement throughout the completion of my research. My profound gratitude also goes to all tutors from Information System Security course, faculty of ACES and Sheffield Hallam University in general for the enhanced quality of the training, support, and services.

References

[1] P. Liu, S. Li, and H. Wang, "Steganography integrated into linear predictive coding for low bit-rate speech codec," Multimedia Tools Appl, vol.76, issue.2, pp.2837-2859, 2017.

[2] G. J. Simmons, "The prisoners problem and the subliminal channel,"Advances in Cryptology, pp.51-67, 1987.

[3] J. E. Storms, "An evaluation of the history, demand, and current methods for digital steganography, " 2016.

[4] T. Sarkar and S. Sanyal, "Reversible and irreversible data hiding technique," 2014.

[5] N. Provos and P. Honeyman, "Hide and seek: An introduction to steganography," Security and Privacy, IEEE, vol.99, no.3, pp.32-44, 2003.

[6] M. Barbier, J. L. Bars, and C. Rosenberger, "Image watermaking with biometric data for copyright protection," 2015.

[7] N. F. Johnson and S. Jajodia, "Exploring steganography: Seeing the unseen," Computer, vol.31, no.2, pp.26-34, 1998.

[8] E. Conrad, S. Misenar, and J. Feldman, Domain 3-chapter 4: Cryptography, 2010.

[9] H. Wang and S. Wang, "Cyber warfare: Steganography vs. steganalysis," Commun ACM, vol.47, no.10, pp.76-82, 2004.

[10] G. Luo, X. M. Sun, L. Y. Xiang, and J. W. Huang, "An evaluation scheme for steganalysis-proof ability of steganographic algorithms,"Intelligent Information Hiding and Multimedia Signal Processing, IIHMSP 2007. Third International Conference on, pp.126-129, 2007.

[11] R. J. Anderson and F. A. P. Petitcolas, "On the limits of steganography," Selected Areas in Communications, IEEE Journal on, vol.16, no.4, pp.474-481, 1998.

[12] S. Kaur, S. Bansal, and R. K. Bansal, "Steganography and classification of image steganography techniques,"Computing for Sustainable Global Development (INDIACom), International Conference on, pp.870-875, 2014 .

[13] J. T. Jackson, G. H. Gunsch, and G. B. Lamont, "Blind steganography detection using a computational immune system: A work in progress," International Journal of Digital Evidence, vol.4(1), pp.19, 2003.

[14] Miller, "Unveiling cyber threats that can impact investigations,", [Online]. Available: http://encase-forensic-blog.guidancesoftware.com/2013/07/c-tak-by-wetstone.html, 2013

[15] H. Kim, N. Bruce, S. Park, and H. Lee, "EnCase forensic technology for decrypting stenography algorithm applied in the PowerPoint file,"2016 18th International Conference on Advanced Communication Technology (ICACT), pp.1-1, 2016.

[16] studentweb, [Online]. Available: http://studentweb.niu.edu/9/ Z172699/Description.html.

[17] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," IBM Systems Journal, vol.35, no.3.4, pp.313-336, 1996.

[18] R. Jafari, D. Ziou, and M. M. Rashidi, "Increasing image compression rate using steganography," Expert Syst.Appl., vol.40, no.17, pp.6918-6927, 2013.

[19] C. Hosmer, "Discovering hidden evidence," Journal of Digital Forensic Practice, vol.1, no.1, pp.47-56, 2006.

[20] A. Cheddad, J. Condell, K. Curran, and P. Mc Kevitt, "Digital image steganography: Survey and analysis of current methods," Signal Process, vol.90, no.3, pp.727-752, 2010.

[21] G. C. Kessler, "An overview of steganography for the computer forensics examiner," Forensic Science Communications, vol.6, no.3, pp.1-27, 2015.

[22] M. S. Sutaone and M. V. Khandare, "Image based steganography using LSB insertion,", no.535, pp.146-151, 2008.

[23] A. Latham, "JPHIDE and JPSEEK stenography programs,", [Online]. Available: http://linux01.gwdg.de/ alatham/stego.html, [Accessed: 26- March- 2016], 1999.

[24] A. Zaharis, A. Martini, T. Tryfonas, C. Illioudis, and G. Pangalos, "Reconstructive steganalysis by source bytes lead digit distribution examination," 2011.

[25] A. Chorein, , [Online]. Available: http://www.silenteye.org/about.html?i6, [Accessed: 27- March- 2016].

[26] K. Magee, "CISSP steganography, an introduction using S-tools,", [Online]. Available: http://resources.infosecinstitute.com/cissp-steganography-an-introduction-using-s-tools/, [Accessed: 29- March- 2016].

[27] EmbeddedSW, "Advanced embedded solutions,", [Online]. Available: http://embeddedsw.net/.

[28] S. Vaidya, "OpenStego, the free steganography solution,", [Online]. Available: http://www.openstego.com/contact.html.

[29] QuickCrypto, "QuickStego,", [Online]. Available: http://www.quickcrypto.com/free-steganography-software.html.