EnCase Forensic Technology for Decrypting Stenography Algorithm applied in the PowerPoint file

HyunHo Kim*, Ndibanje Bruce*, SuHyun Park**, HoonJae Lee**

*Dept of Ubiquitous IT, Dongseo University, Korea

** Dept of Computer Engineering, Dongseo University, Korea

feei_@naver.com, ndibanje.bruce.phd@ieee.org, subak@dongseo.ac.kr, hjlee@dongseo.ac.kr

Abstract— **The growth of both IT technology and the Internet Communication has involved the development of lot of encrypted information. Among others techniques of message hiding, stenography is one them but more suspicious as no one cannot see the secret message. As we always use the MS Office, there are many ways to hide secret messages by using PowerPoint as normal file. In this paper, we propose a new technique to find a hidden message by analysing the in PowerPoint file using EnCase Transcript. The result analysis shows that Steganography technique had hidden a certain number of message which are invisible to naked eye.**

Keywords— **Encase, Forensic, Stenography, PowerPoint**

I. Introduction

Development of Internet and IT technology have opened a lot applications to facilitate the information transaction via Internet using computers and mobile devices. With the ubiquitous features, a wide range of information can be accessed and thus, information security is very important. In addition advanced techniques in hiding information technology are becoming diverse, and representative technology is a steganography. Steganography technique is mainly based on hiding a text, audio, and image, in addition of movie [1], embedded in a chosen file. Usually it is hard to figure out hidden message if not you need to use the same application that was used to encrypt the hidden message.

Nowadays, secret message can be hidden using program function and can be sent to the destination (people, or elsewhere). In accordance with examples of using MS Office PowerPoint overlaps the picture and leave a message hidden in the picture, place or by modulating the hyperlinks or the sites that are connected to different physical, font colour, etc., using compulsory to hide. In this paper, using the EnCase we analyse PowerPoint MS Office file in order to find out hidden message by stenography techniques

II. Related work

For investigation purpose and professional work, it is very important to use digital forensic tool instead of an application data. For instance the data is the name of the person who created the document, the contents may be harmful to the user, such as a summary of the document and may be immoral or illegal. In case of forensic investigator all data should be extracted and can be analysed. If any harmful is caught then the investigation should proceed. According to Byers, hidden message was found in each file of 100,000 MS Word collected document files at random

A. Previous research(MS Office 2003, 2007)

The official name of PowerPoint (MS) was micro post Office PowerPoint. PowerPoint allows the information based on the product, service, integrated solutions to one effective work [2]. In 2007, A. Castiglione et al. [3] presented a research report on digital forensic and steganography where they mentioned that earlier versions of MS Office 2003 can hide message. Furthermore, B. Park et al. revealed that MS Office 2007 can hide message [4].

Nowadays, the hiding techniques from MS Office has evolved because, program function can be used to hide message such as: text points, image overlaps, the image size, font color, background color, etc.) and also it has been known that it not easy to find out that hidden message.

B. Steganography

Steganography algorithm one of the art of hiding a message, has come back to the researchers attention during the past few years as it can be a very efficient way to increase the security of a system. It can be used both on its own, in systems which can benefit from hiding data but do not need encryption or other forms of security, and combined with another security method. The most used combination is that of steganography and cryptography. Both methods can be sufficient used separately, each according to what is needed in a certain context. However, sometimes, complex systems might benefit from the usage of two such methods. For example, an encrypted message might be seen by a third party and even though without the proper keys, decryption would be hard, someone would still know we are sending encrypted messages. In many cases this could be considered a breach in security.

However, hiding the encrypted message within a text file, or an image could increase the chances of that message never being intercepted. Thus, the usage of steganography comes in useful. Another scenario in which we can add steganography is to hide the keys being sent to users - either the private keys they will use for decryption or those used for digital signatures, which is also the case in the present paper. These algorithms can be very secure from a computational point of view and once the message is encrypted, or signed respectively, decoding it becomes almost impossible. However, if the keys are intercepted the security of the system becomes compromised.

In the digital signature algorithm we can see two situations in which steganography can be useful. For both situations we used image steganography with the least significant bit (LSB) method[4-7]. It is probably the best known steganography method as it covers its requirements perfectly: to hide data so that the changes in the original image are not visible to the naked eye. The method is also rather easy to use as it does not take a lot of processing time, it does not slow down the system. For all these reasons we considered the least significant bit method to be perfect for our system. The digital signature module is used for role definition and access control and the algorithm used is robust and safe even when used on its own. For this reason, we needed to hide the keys in a way that is efficient, yet simple, so as to not overload the system with this operation.

C. *Transcript*

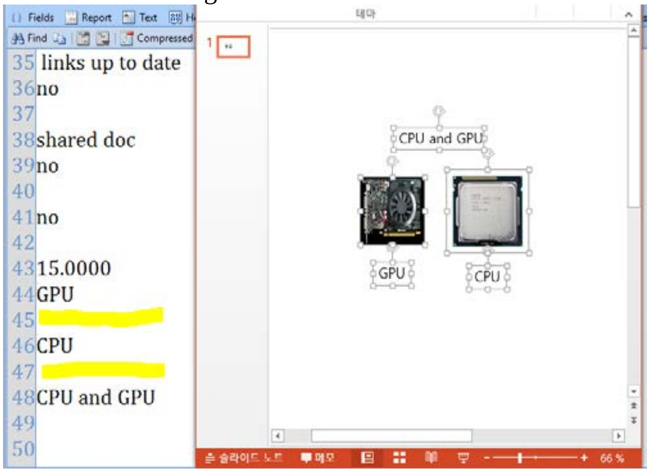
The Transcript view is also an integration of Stellent Inc.'s Outside In Technology. This view suppresses file noise, such as formatting and metadata. The text displayed within this view is text that is typically indexed for search by the indexing engine. Any search hit or bookmark will appear in both the Doc and Transcript view[8].

III. Analysis of Information Hiding Type

In this section, we describe how to find out the type of information hidden in PowerPoint. In this experiment analysis, we have used programs such as EnCase v7 and MS Office2013 and we use default function from PowerPoint during information and the analysis of hidden information has been done by EnCase Transcript.

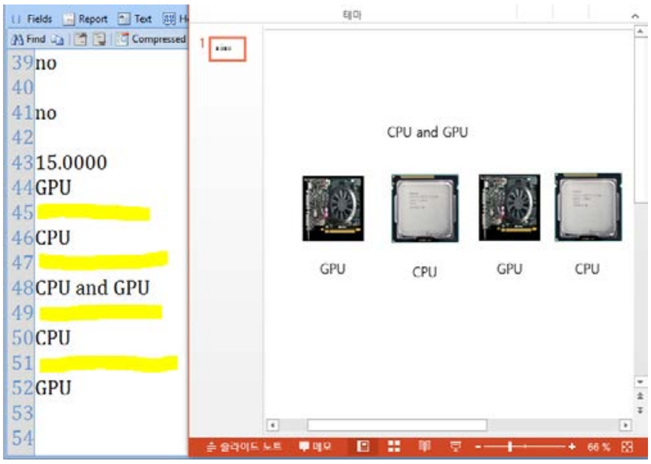
A. *Typical analyses : images and text boxes*

In this subsection, the images and text boxes shown in Figure 1 and Figure 2. describe how the EnCase Transcript finds out the images and results in the PowerPoint slides.



```
35 links up to date
36no
37
38shared doc
39no
40
41no
42
4315.0000
44GPU
45
46CPU
47
48CPU and GPU
49
50
```

Figure 1. 2 Image Transcript



```
39no
40
41no
42
4315.0000
44GPU
45
46CPU
47
48CPU and GPU
49
50CPU
51
52GPU
53
54
```

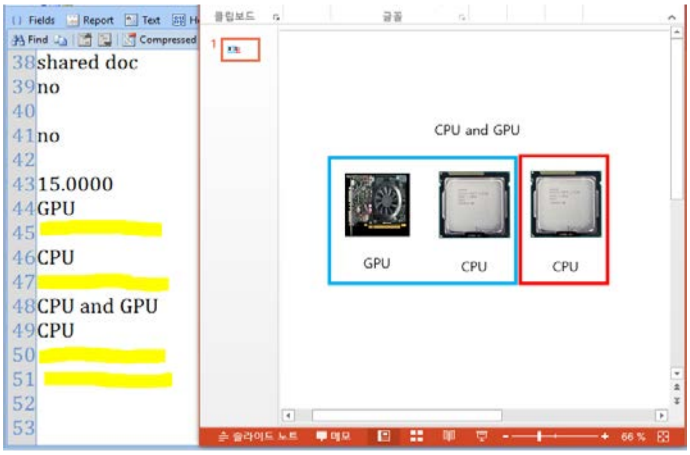
Figure 2. 4 Image Transcript

In Figure 1 two images are in the right side, and three text box. Figure 2 shows an image consisting of four images and five text boxes. In addition, in Figure 1, the text box (CPU and GPU) are observed in Transcript on left side displayed on line 48. Similarly, Transcript each image is displayed in line 45 and 47(yellow colour: 45, 47). Figure 2, two additional text box and two images are displayed, Transcipt on left side has added also the lines (49-52).

B. *Analysis of overlap image*

This case describes a hidden image in the red box where in Figure 3 below, the left (blue box) showed the original image but the images in the right (red box) has been overlapped.

From the Transcript analysis in left side, it is reported that there is four images instead of three as displayed in slide. The yellow lines (45, 47, 50, 51) has been confirmed that the four images by Transcript analysis. As result, if any image is added to the slide, Transcript also will add a more line in left side. Consequently, it is visible from the Transcript analysis that there is one hidden image.

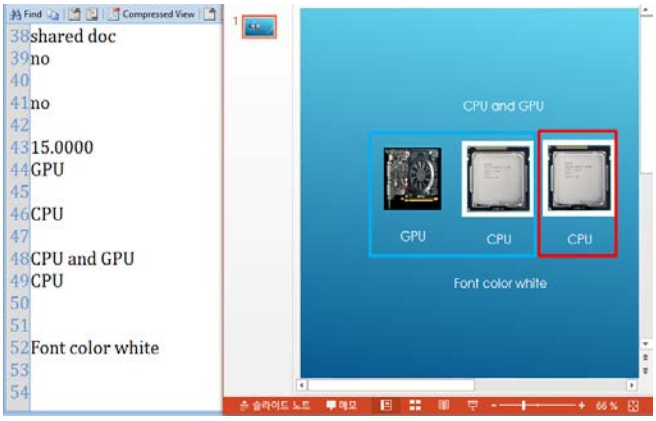


```
38shared doc
39no
40
41no
42
4315.0000
44GPU
45
46CPU
47
48CPU and GPU
49CPU
50
51
52
53
```

Figure 3. Normal image (blue) and overlap image (red)

C. Similar Background colour with Text Box Color

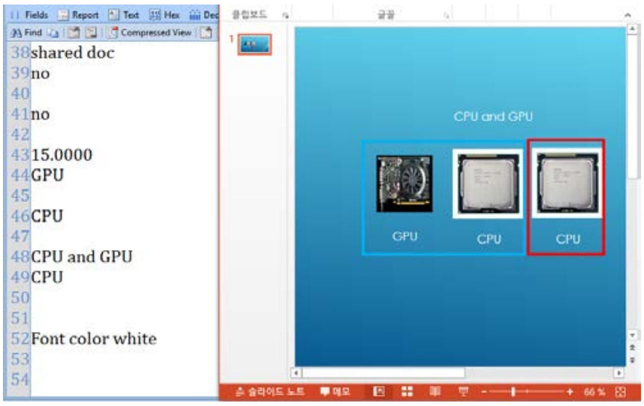
In this subsection, we describe a hidden message by changing background colour to the text box. Before changing the background colour Transcript can display the text box message in the left side as shown if Figure 4 .Therefore, the colour is changed in Figure 5 and result analysis is given by Transcript.



```
38shared doc
39no
40
41no
42
4315.0000
44GPU
45
46CPU
47
48CPU and GPU
49CPU
50
51
52Font color white
53
54
```

Figure 4. Transcript Analysis: “Font color white" in left side

In Figure 4 we can observed the message from the text box message is: "Font color white" and Transcript is displaying the same message in the left side to the line 52.



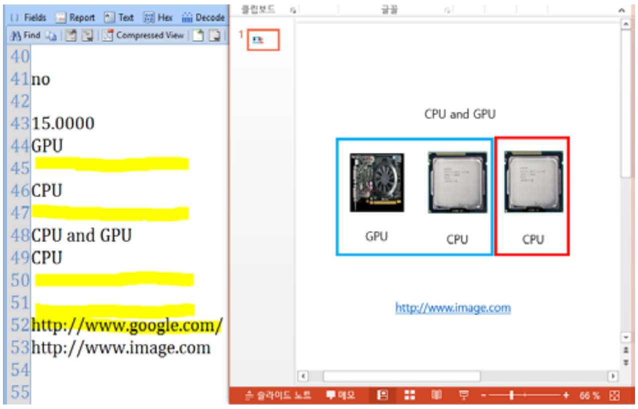
```
38shared doc
39no
40
41no
42
4315.0000
44GPU
45
46CPU
47
48CPU and GPU
49CPU
50
51
52Font color white
53
54
```

Figure 5. After changing text color

In Figure 5 the text box message is the same as in Figure 4: "Font color white" but the color is changed to the similar background color. As result we have a hidden message in Figure 5. In the view of a naked eye, it is not impossible to see the hidden message as described we can only observe 3 imaged and 4 text boxes. To find the hidden message in this particularly case is really very hard. Therefore, it now time to require Transcript analysis in order to discover the hidden message.

D. Fake hyperlinks

Figure 6 shows how to fake hyperlink. As described in Figure 4 and 5 the method of hidden information from text box by changing background color can be applied to hyperlink embedded into the slide. Normally, hyperlink function is to direct to the website in order to get more information. But for hidden purpose someone can change from real website to wrong one. In case of Figure 6 the displayed hyperlink in slide shows: "http://www.image.com" but in reality the direction of to the link is "http://www.google.com".



```
40
41no
42
4315.0000
44GPU
45
46CPU
47
48CPU and GPU
49CPU
50
51
52http://www.google.com/
53http://www.image.com
54
55
```

Figure 6. Transcript : hyperlink

Therefore, Transcript analysis is reporting different view as we can observe in the left side. The link address displayed in the slide is also seen by Transcript (line 53: http://www.image.com), we can see a second Link Address

(Line 52: http://www.google.com) which is not displayed in the slide of the right side.

As result Transcript proof that there is a hidden message which can be a text box or hyperlink address that has been changed or faked.

IV. CONCLUSIONS

In this paper, we conducted a research where analysed types of hidden information in the PowerPoint using the EnCase find out hidden message. In addition techniques of hiding information using PowerPoint were described and comparison result from the Transcript analysis showed that steganography techniques has been utilized to hide the information in the PowerPoint file.

Furthermore, we have analyse a case of information hiding type using hyperlink where the displayed hyperlink does not direct the real website. Also we have discussed about overlap images, change of background color in regards with text box message.

Finally, the results from all experiment have revealed that nowadays steganography techniques can be applied to PowerPoint file in order to hide messages but Transcript analysis always has been useful tool to detect and discover the hidden message. In the near future, the PowerPoint will be very useful for digital forensic.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology. (grant number: NRF-2011-0023076).

REFERENCES

[1] Seon-su Ji., A Study and improved Approach of Text Steganography, *Journal of the Korea Industrial Information Systems Research.*, Vol. 19 No.5, Oct. 2014.

[2] http://terms.naver.com/entry.nhn?docld=1224889&cid=40942&categoryId=32837

[3] A. Castiglione, De Santis, C. Soriente., "Talking advantages of a disadvantage : Digital forensics and steganography using document metadata", The Journal of Systems and software, vol 80, Issue 5, pp.750-764, 2007.

[4] Bora Park, JungHeum Park, Sangjin Lee., “Information Hiding and Detection in MS Office 2007 file”, Journal of The Korea Institute of Information Security and Cryptology Vol 18, No.3, June, 2008.

[5] A. H. Ibrahim, W. M. Ibrahim., “Text Hidden in Picture using Steganography: Algorithm and Implications for Phase Embedding and Extraction Time”, International Journal of Information Technology and Computer Science, Vol. 7, No. 3, Jan/Feb 2013.

[6] Riadh W. Y. Habash, Voicu Groza, Kevin Burr., “Risk Management for Power Grid Cyber-Physical Security, British Journal of Applied Science and Technology, Vol. 3, Issue. 4, July 2013.

[7] G. Padmashree, P. S. Venupogapala., “Audio Steganography and Cryptography: using LSB algorithm at 4th and 5th LSB layers”, International Journal of Engineering and Innovative Technology, Vol. 2, Issue. 4, Oct 2012.

[8] Steve Bunting, EnCE., “EnCE The Official EnCase Certified Examiner STUDY GUIDE Second Edition” 2008.

HyunHo Kim

2013 : BS at Dongseo University, Republic of Korea
2015 : MS at Dongseo University, Republic of Korea
2015 ~ current: doctor´s course Dongseo University, Republic of Korea
Research Interests : Digital Forensic, Information Security, Network Security

Ndibanje Bruce
2004 : BS at Ngozi University, Republic of Burundi
2013 : MS at Dongseo University, Republic of Korea
2016 : Ph.D at Dongseo University, Republic of Korea
Research Interest: Information Security, Wireless Sensor Networks, Cryptography and Network Security, Side Channel Analysis

SuHyun Park
1986 : BS at Pusan National University, Republic of Korea
1988 : MS at Pusan National University, Republic of Korea
1999 : Ph.D at Pusan National University, Republic of Korea
1996 ~ current : Professor of Dongseo University, Republic of Korea
Research Interests : Maritime IT, Artificial Intelli

gence, Intelligent System

HoonJae Lee
1985 : BS at Kyungpook National University, Republic of Korea
1987 : MS at Kyungpook National University, Republic of Korea
1998 : Ph.D at Kyungpook National University, Republic of Korea
2002 ~ current : Professor of Dongseo University, Republic of Korea
Research Interests : Password Theory, Network Sec

urity, Side-Channel Attack, Information Communication/Information Networ

k