Journal of the Korea Institute of Information Security & Cryptology

Volume 28 Issue 6 / Pages.1439-1448 / 2018 / 1598-3986(pISSN) / 2288-2715(eISSN)

Korea Institute of Information Security and Cryptology

Study on History Tracking Technique of the Document File through RSID Analysis in MS Word

Research on document file history tracking technique through RSID analysis of MS Word

DOI QR Code

Joun, Jihun (Institute of Cyber Security & Privacy (ICSP), Korea University) ; Han, Jaehyeok (Institute of Cyber Security & Privacy (ICSP), Korea University) ; Jung, Doowon (Institute of Cyber Security & Privacy (ICSP), Korea University) ; Lee, Sangjin (Institute of Cyber Security & Privacy (ICSP), Korea University) Jeon Ji-Hoon (Korea University Graduate School of Information Security) ; Jaehyuk Han (Korea University Graduate School of Information Security) ; Doowon Jeong (Korea University Graduate School of Information Security) ; Sangjin Lee (Korea University Graduate School of Information Security)

Received : 2018.08.21 Accepted : 2018.11.14 Published : 2018.12.31

https://doi.org/10.13089/JKIISC.2018.28.6.1439 Copy Citation KSCI HTML

Abstract

Many electronic document files, including Microsoft Office Word (MS Word), have become a major issue in various legal disputes such as privacy, contract forgery, and trade secret leakage. The internal metadata of OOXML (Office Open XML) format, which is used since MS Word 2007, stores the unique Revision Identifier (RSID). The RSID is a distinct value assigned to a corresponding word, sentence, or paragraph that has been created/modified/deleted after a document is saved. Also, document history, such as addition/correction/deletion of contents or the order of creation, can be tracked using the RSID. In this paper, we propose a methodology to investigate discrimination between the original document and copy as well as possible document file leakage by utilizing the changes of the RSID according to the user's behavior.

Various electronic document files, including MS Word, have become a major issue in various legal disputes, such as contract forgery and trade secret leakage. A unique RSID (Revision Identifier) is stored in the internal metadata of files in the OOXML (Office Open XML) format used since MS Word 2007. RSID is a unique value assigned to a word, sentence, or paragraph every time the content of a document is created/edited/deleted and then saved. It records document history such as content addition/editing/deletion history, creation order, and document application used. It can be estimated. In this paper, we present a methodology to investigate changes in RSID according to user behavior, distinguishing originals from copies, and leaking document files.

Keywords

Revision Identifier ; Document forensics ; OOXML ; MS Word

I. Introduction

Unlike paper documents, electronic documents can be easily modified and copied, which can lead to various illegal activities such as document leakage, copyright infringement, plagiarism, and contract forgery. In addition, electronic documents are one of the important analysis targets in digital forensics because malicious programs can be inserted into document files or confidential information can be hidden to communicate with each other [1].

A representative prior study that distinguishes between originals and copies is a method of investigating fraud by comparing attribute information [2]. However, methods for falsifying attribute information are known, and several tools for changing time values already exist, allowing for easy falsification. Using the RSID (Revision Identifier) presented in this paper, it is possible to check whether there is an anti-forensic activity that modifies the time value, number of modifications, modification time, etc.

RSID is a unique value inside a document file, and without using other attribute information, it is possible to know whether the document itself or some contents within the document have been copied, and the file creation order and history can be known. The file creation history shows the document editing program used by the author when creating the document file and the source of the written content (whether copied from another MS Word document or another application or written directly). By comparing only the RSID, you can quickly find only relevant document files from a large amount of data, and you can protect your personal information by conducting research without viewing the document. In addition, RSID is used not only in MS Word but also in various document files. In particular, in LibreOffice, the basic document program in Linux, RSID is generated according to rules similar to MS Word, so it is possible to know whether it has been copied, the history of the document file, etc., enabling document file forensics. It can be mainly used in.

II. Related research and background knowledge

MS Office versions after MS Office 2007 use OOXML as the default format. The structure of OOXML, which includes attribute information and the RSID used in this paper, is shown in Fig. Same as 1. The core.xml and app.xml files in the docProps folder store metadata such as creation time, modification time, author, and number of modifications of the file. However, this attribute information can be easily changed. There are various types of RSIDs in the "documents.xml" and "settings.xml" files in the folder [3]. Through this, you can find out whether the content has changed or whether it has been copied.

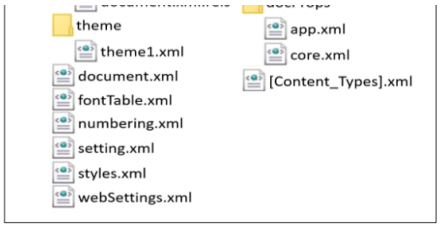

Fig. 1. OOXML Structure

Fig. 1. OOXML Structure

Fu Z [2] presented two forensic analysis methods targeting the OOXML format. The first is to check the contents of the MS Word file and internal attribute information (file size, creation date, modification date, access date, creator, last saver, number of modifications) of the MS Word file, which is a method already used. . However, there is a limitation that it can be abused because attribute information can be easily changed using tools or by directly modifying it. The second method is a forensic analysis method using RSID. When copying files, the author proposes a method to check whether the two files are copied using rsidR and rsidRPr in the "document.xml" file and RSID in the "settings.xml" file. did. However, there is no description of the types and creation rules of other RSIDs other than rsidR and rsidRPr, so the creation history cannot be known. In addition, when MS Word document files are created using Hancom Office, one or more RSIDs always appear the same, and there is a possibility of false positives in the method proposed by the author. Therefore, in this paper, a complementary MS Word file forensic analysis method is used. present.

Espen Didriksen [4] studied forensic methods targeting the OOXML format in more depth. In addition to the analysis method using attribute information, the author explained the definition of RSID and definitions by type and presented various forensic analysis methods. However, the author's paper does not contain a detailed explanation of the creation rules for each RSID type, making it difficult to investigate file history. However, in this paper, we organize RSID generation rules according to user behavior, explain specific copy history and file editing history analysis methods, and present efficient investigation procedures for practical use.

In a paper [5] that analyzed a real case using RSID, the terrorist manual distributed on the day of the attacks in Oslo and Uttuya, Norway, was reviewed to see if it was consistent with the claims of the suspect arrested for terrorist acts, and another author Analyzed whether it exists. The author of the paper used RSID to confirm that the terrorist manual was saved from another source. The author conducted the analysis using only rsidR among the RSIDs. rsidR, rsidRDefault, rsidRPr, and rsidP, which will be explained in this paper, were used to check which content was copied and stored from other sources, and to track the creation history to obtain more detailed results. can be obtained.

Additionally, there is a method to block malicious code by collecting document files containing hidden malicious code in a database and checking the RSID before viewing files received from outside [6]. Using RSID not only has the advantage of not being infected by malicious code, but also has the advantage of being able to find relationships without viewing the contents of the file when examining document files.

In addition to MS Word files, you can also check the writing history in PDF files and MS Excel files. Hyunji Chung [7] presented a method to check remaining data even after modifying a PDF file. He explained that the creation history of a PDF file can be checked by tracking data before modification, and that this area can also be used to hide data. Yoon Mi and Lee [8] explained forensic investigation methods and a method of tracking writing order using MS Excel metadata.

There is a method for analyzing the creation history of MS Word files using temporary files [9], but it has the limitation that it is only possible if a temporary file exists, so if the file that can be analyzed is a single file, the investigation is not possible. It's difficult, but using RSID, you can find out the file history with just one MS Word file.

As studies estimating file history by document file type, such as the papers mentioned above, are continuously emerging, document file forensics is useful in criminal investigations. However, while conducting an actual document file forensics case, there are cases where metadata such as time value or number of modifications have been intentionally falsified, causing difficulties in file forensic analysis. To solve this problem, the RSID presented in this paper can be used to prove that the file has been altered. In addition, existing papers related to RSID can only be used for copy history between two files, but this paper presents various utilization methods using RSID mentioned in the introduction.

III. RSID (Revision Identifier)

RSID is a 32-bit unique value assigned inside the file every time the document contents are changed and saved. Among the 8 digits of RSID, the first two digits are fixed to "00", and the last six digits are composed of random hexadecimal numbers. In theory, the probability that RSIDs randomly overlap with the same value is $1/16^{6} = 1 / 16,777,216$. As a result of comparing the RSIDs of 1,000 files obtained by crawling random docx files from the Internet with a tool developed in Python, it was confirmed that there were 716,221 RSIDs, and there were no duplicate RSIDs.

3.1 Comparison of differences between RSID types

As shown in Table 1, RSID is divided into seven types depending on the range of elements that make up the text, such as sentences, paragraphs, tables, and sections. The types of this are defined in existing papers [3], but additional research is needed on the modification range and generation rules that each RSID represents.

Table. 1.Types of RSID

Table. 1. Types of RSID

Type	Paragraph	Run	Table	Section
rsidR	w:p w:rsidR	w:r w:rsidR	w:t w:rsidR	w:sectPr w:rsidR
rsidRDfault	-	w:r w:rsidRDfault	-	-
rsidRPr	w:p w:rsidRPr	w:r w:rsidRPr	w:t w:rsidRPr	w:sectPr w:rsidRPr
rsidDel	w:p w:rsidDel	w:r w:rsidDel	w:t w:rsidDel	w:sectPr w:rsidDel
rsidP	w:p w:raidP	-	-	-
rsidTr	-	-	w:t w:rsidTr	-
rsidSect	-	-	-	w:sectPr w:rsidSect

rsidR is created by default when a document is first created, and is subsequently created when a new word, sentence, or paragraph is created within the document. In other words, it is created by writing or copying new content following an existing sentence.

rsidRDefault is created by default when a document is first created, and is created only when a new paragraph is created. In other words, if you create a document with a line break or copy a paragraph containing a line break, it is generated along with rsidR.

rsidRPr is created when a change occurs in the font. When the font is changed, it can be roughly divided into two types. The first is when the font of a text that has already been written is changed, and the second is when the font of a text in another file is different from the file and the text is copied and pasted. . In other words, it is created when changing the font of existing content or copying content written in a document application other than MS Word (not the font of the file). Fig. If you change "Forensics" to Arial Black font in "Digital Forensics" as shown in 2, a new rsidRPr is assigned to the "document.xml" file and the corresponding RSID is also assigned to the "settings.xml" file.

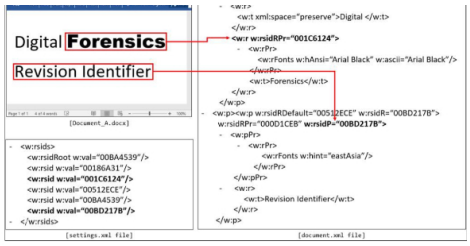

Fig. 2. document.xml and settings.xml in RSID_document.docx

Fig. 2. document.xml and settings.xml in RSID_document.docx

rsidP is created when you copy a new paragraph in MS Word or another document application. A new paragraph is a paragraph that contains a line break (enter). Fig. If you copy and paste the "Revision Identifier" from another MS Word file as shown in 2, a new rsidP is assigned to the "document.xml" file and the corresponding RSID is also assigned to the "settings.xml" file.

rsidDel is an RSID that is generated when an entire paragraph is deleted from the content of the text. However, deleting a paragraph does not always create it, but only creates it in certain cases, and this part requires further research. rsidTr is an RSID created when the table is modified, and rsidSect is an RSID created when the layout (paper size, paper direction, margin, etc.) is changed. Espen Didriksen[4]'s paper explained that the RSID does not exist after MS Word 2010, but it was confirmed that it exists even in the latest version.

Next, in order to estimate the history when editing a file, the RSID generation method is explained in the following order.

1) Copy the original file

2) Add/edit/delete contents of copied file

3) Copy the contents of the original file

4) RSID characteristics by document application

3.2 Experimental results to identify RSID characteristics

This paper was conducted with MS Word 2016 on Windows 10. RSID is created according to the same rules in MS Word 2003, 2007, 2010, 2013, and 2016 without any major differences. In addition, RSID is assigned according to the same rules even when content is created/edited/deleted in Windows XP, Vista, 7, 8, 10, and MAC OS, and the RSID does not change even if the MS Word file is moved by operating system.

① When copying a new MS Word file, there are two ways to copy an MS Word file: 1) "copy-paste" the original file itself, and 2) use the "Save As" function in the original file. . When copying files using method 1), the RSIDs all match the original files. However, when method 2) is used, one RSID is created additionally to the RSID of the original file, and its value is recorded in "settings.xml". You can find out how the file was copied by comparing the list of RSIDs of the original file and the RSID of the copied file.

② When adding/editing/deleting content in a copied MS Word file, create a sentence in a new document, or add content to an existing sentence and save it, the new rsidR and rsidRDefault will be set to "document.xml" in the added location. and the corresponding RSID is added to the "settings.xml" file. When editing/deleting document content, if at least one character remains among the content that has already been assigned an RSID, the existing RSID remains in the "document.xml" file. However, if you delete all the contents, all existing RSIDs will disappear. However, even if all RSIDs in the "document.xml" file disappear, all RSIDs once saved in the "settings.xml" file remain. Additionally, a new RSID created upon deletion is additionally recorded in "settings.xml".

By using the RSID rule that changes when created/edited/deleted, when there is an original, a copy, or an MS Word file suspected to be a leaked file, even if the leaker modifies the file or deletes the contents, the existing file in the "settings.xml" file Since all RSIDs are recorded, the relationship between the file and the original file can be proven.

③ When copying the contents of an MS Word file to another document, before explaining copying the contents, the basic font mentioned in this paper refers to changing the font size, color, thickness, italics, etc. when creating a file in MS Word. It was written without . In other words, it is written in Clear Gothic font in the Korean version and "Calibri" font in the English version. If the content to be copied is a default font, if you copy the content written in the default font and paste it into a new document or another document, the RSID assigned to the existing file will not be copied to the copied file. In other words, even if content written in the default font without changing the font and style is copied, a new RSID is assigned to the copied content. If the font of the content to be copied is not the default font, unlike when copying the default font, if the font is modified, a new RSID called rsidRPr is assigned to the modified content, and when copying to a new document, the original rsidRPr is copied. . When the font of part of the document content is changed (size, style, color, font type), the original content containing at least one letter among the changed content is copied to the new document, and the RSID is copied and saved as is. Depending on the character style, rsidRPr may or may not be created, which is summarized in Table 2. However, when copying and pasting from another MS Word file, if you select the "Keep text only" option rather than the "Keep original formatting" option, rsidRPr is not created and a new value different from the original value is assigned to the RSID.

Table 2. Types of styles that create rsidRPr

Table 2. Types of style that create rsidRPr

Type	rsidRPr created
Bold/Italic	O
Array	X
Color	O
Font Type	O
Font Size	O
Listing	X
Style	O
Chart	X
Picture	X
Text highlight	O

3.3 Comparison of RSID analysis results by document editor

Users may not write only in MS Word, but may write using other document editors or modify in other operating systems. This subsection explains changes made when an MS Word file is created and modified in another word editor.

RSID is not only created in MS Word, but may also be created in other document editors. There are programs created with the same rules, but in some cases they only exist and cannot be used digitally forensically. Table 3 shows the presence or absence of RSID for each document editor. The type of editor used by the user can be

estimated through the RSID in the document created/edited with each document editor.

Table 3. Existence of RSID according to application

Table 3. Existence of RSID according to application

Application type	RSID existed
MS Word	O
MS Office Online	O
Office 365	O
Google Document	O
Mobile MS Office	O
Libre Office	O
Open Office	X
Naver Office	X
Hancom Office	O

In the case of MS Office Online Word, when a new document file is created, a "document2.xml" file is created instead of "document.xml" in the document, and the RSID is present in it. In existing MS Word, the first two 8 bits are fixed to "00" and the last 6 digits, 24 bits, are randomly generated, but when a new document is created in MS Office Online Word, the first two 8 bits are set to "5A4C71E5". A value other than "00" is assigned. When modified with MS Word after being created in MS Office Online Word, the "document2.xml" file is changed back to "document.xml" file, and the modified part generates the same RSID as the existing MS Word. Other than that, RSID creation rules are the same as MS Word. Therefore, if "document.xml" does not exist in the file but "document2.xml" exists, or if the first two 8 bits of the RSID are assigned to a hexadecimal number other than "00", it is first created/created in MS Office Online Word. It has been modified.

In files created with Office 365, RSIDs are created according to the same rules as MS Word, but two more RSIDs are created in the "settings.xml" file that do not exist in the "document.xml" file. Hancom Office has a fixed value, but Office 365 has the characteristic of changing randomly.

When you create a new Google document or save a document created in MS Word in Google Document, all RSIDs are assigned "00000000". Therefore, if at least one RSID in the file is "00000000", it was created/edited using Google Document.

MS Office for mobile generates RSIDs according to the same rules as existing MS Word. Even if you move a document file created in MS Office for iOS to Windows OS, the RSID does not change at all, and is created according to the same rules as when editing a document in MS Word.

In the case of libre office, the RSID exists in the "content.xml" file rather than the "document.xml" file and also exists in the "settings.xml" file, but the RSID creation rules are different from MS Word. When editing a document using Libre Office, RSIDs are assigned to the tags "paragraph-rsid" and "officeooo:rsid" in "content.xml", and words and paragraphs containing "officeooo:rsid" When copied to another document, it is always copied and saved regardless of whether the font has been changed.

It was confirmed through experiments that RSID is not created in Open Office and Naver Office, and that RSID disappears when an already created MS Word file is modified.

When Hancom Office 2018 is installed as the default environment, "Hancom Office 2018 Hanword" is created. When you right-click to create a Word file, it is automatically created as Hancom Office Word. This action is installed as a default setting when Hancom Office 2018 is installed, so when most users right-click to create a Word file, it is written in Hancom Office Word. In this case, the RSID rules are different from when writing in existing MS Word. When writing with Hancom Office, the first rsidR in the "document.xml" file is always assigned a fixed value. In Hancom 2018 version, the first rsidR is fixed to "002764DB", and rsidRoot in the "settings.xml" file is fixed to "00506824", and this value is repeated twice and saved. This value remains even if all strings are deleted, so if the fixed RSID above is present in the file, you can know that the document was created using Hancom Office. Therefore, when comparing the RSIDs of document files to find relationships, "002764DB" and "00506824" must be excluded.

IV. Document file investigation plan through RSID analysis

You can check leaks or distinguish between originals and copies through RSID analysis in MS Word. In particular, it is possible to determine whether there is plagiarism because the writing order of the contents in the file can be identified. There are various analysis methods using RSID, but it has the limitation that it can be falsified like other metadata. However, if the RSID of "document.xml" and "settings.xml" is tampered with, the time value changes from "1980-01-01 12:00:00 AM" to the current time, allowing you to check for tampering [4]. Additionally, unlike existing metadata research methods, there is no tool to change the RSID yet, and in order to tamper with the RSID without a trace, one must know the rules of the RSID described in this paper. According to the RSID rules specified in existing RSID-related papers, if the RSID in a file is arbitrarily deleted or changed, traces of tampering that violate the RSID rules can be found.

The first step for analysis is to extract rsidR, rsidRDefault, rsidP, and rsidRPr from the "document. do. Fig. 3 is a diagram schematizing the process.

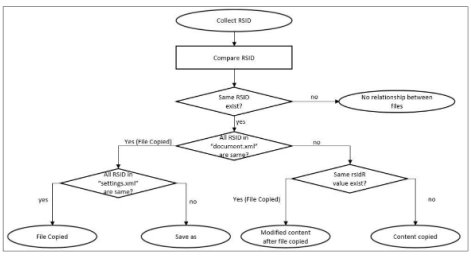

Fig. 4. Advanced investigation process of document file leaking or plagiarism

Fig. 3. Basic investigation process of document file leaking or plagiarism

4.1 Investigation of external leakage and plagiarism of document files

In order to investigate whether a MS Word file has been leaked, the RSIDs present in the "document.xml" and "settings.xml" files must be distinguished and compared. If you extract all RSIDs of the original document, extract the RSIDs from the MS Word files in the storage medium you want to search, and search for files with the same value first, you can check them accurately and in a short time. Fig. Once the basic research process is completed as shown in 3, the in-depth research process is conducted. The advanced process is as described below, and is shown in Fig. 4 is a schematic drawing.

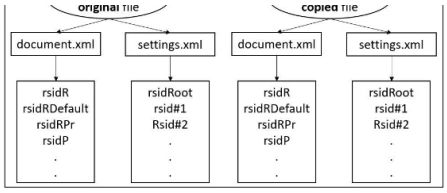

Fig. 3. Basic investigation process of document file leaking or plagiarism

Fig. 4. Advanced investigation process of document file leaking or plagiarism

Compare the RSID of each file, and if there is no same RSID, the two files have no relationship. If the same RSID exists, check whether all RSIDs in "document.xml" of the two files are the same. If all RSIDs in "document.xml" and "settings.xml" are the same, the original file has been copied. If one more RSID is added to "settings.xml", this is a file that saved the original file under a different name. If some of the RSIDs in "document. However, if the same rsidR does not exist in "document. If the rsidR and rsidRDefault of two files are the same, the file with fewer RSIDs in the "settings.xml" file is the original file, and many files are copy files.

4.2 Investigation of document file content modification history

Using the RSID and metadata within the document file, you can track the creation order of the document file. Fig. The "Case_Document.docx" file in 5 is a document file with contents written in a random order, and the writing order cannot be determined just by looking at the contents. In this case, using RSID can reveal the user's document creation order and various information.

JBBHCB_2018_v28n6_1439_f0005.png image

Fig. 5. document.xml and settings.xml in Case_Document.docx

First, the rsidRoot of the "settings.xml" file is "00B45129" and this value is the same as the rsidRDefault of the "document.xml" file, so you can see that it is a file created with "New" in MS Word. Therefore, the author first wrote the word "South Korea" and saved it after a line break. The reason is that the rsidR assigned to the sentence "Seoul Gangnam Station 5pm" in the "document. If you write it at the line change, the rsidRDefault of the first sentence will remain in the next character rsidR, as follows. However, the words "Seoul Gangnam Station 5pm" are in the last act, not the second. In the sentence "Seoul Gangnam Station 5pm"If you look inside , It was confirmed that the tag remains in the last created/edited part. Therefore, this sentence was written last.

The word created in the second action is "Digital Forensics." If you look at the tag containing the word, rsidRPr has a different value for the word "Digital". This means that after writing and saving the word "Digital Forensics," the font for the word "Digital" was modified to red. Able to know. And as mentioned above, the author also wrote "Seoul" in the second line of the last action and copied and saved the sentence "Gangnam Station 5pm." If you check the RSID of the sentence "Gangnam Station 5pm", you can see that rsidRPr has been assigned, so it was copied from somewhere else. If the tag surrounding the sentence contains , If the tag created when the font is changed does not exist, you can tell that rsidRPr was not created because the font changed, but was copied from somewhere else. To summarize the order of actions in this file, "South Korea" was written in the first line, changed to a line and saved, and after changing the line, "Digital Forensics" was written in the third line and saved. Afterwards, I changed "Digital" to red in "Digital Forensics" and saved it. After writing "Seoul" in the second sentence, I copied and saved "Gangnam Station 5pm" from somewhere else.

4.3 Introduction to application cases: Identifying the time of creation of Internet postings and document contents of files subject to analysis

The forensic investigation team secured "File A," which is suspected of being written as an act of plagiarism, and "Post B," which was written on an internet blog and is suspected of being the original file. The suspect claims that he wrote File A himself, and when checking the time value, he claims that "Post B" was written earlier than the date it was uploaded to the Internet. "Post B" was written on May 10, 2018, and the creation time, modified time, and access time of file A are all set to March 5, 2018, and the number of modifications is also 0.

Investigators used RSID to find the suspect's anti-forensic activities and investigated plagiarism. First, there are a total of 3 RSIDs stored in the "document.xml" file of file A. This is different from the number of modifications of "File A", as you can see that it was modified and saved at least three times. Second, ""File A" and "Post B" have many parts of the same content, and the content of the file posted on the Internet in File A has been slightly modified or added. The presence of rsidP in the "document.

In conclusion, contrary to the suspect's claim and the information in "File A," the suspect copied and modified part of the Internet posting, and also changed the time value to the past and falsified the number of modifications through an investigation method using RSID. I found out.

V. Conclusion

In this paper, we studied digital forensic analysis techniques using RSID recorded in MS Word files during document work. RSID is a unique value generated as a document is created, and can be used to determine not only whether the document itself has been copied, but also whether the content within the document has been modified or copied. It is expected that it can be used in various investigation methods such as anti-forensic activities, plagiarism, leakage, and integrity by tracking the writing history, and it is necessary to collect RSID rules to utilize the analyzed results.

* This paper was a study conducted with the support of the Information and Communication Technology Promotion Center with funding from the government (Ministry of Science and ICT) in 2018 (No. 2018-0-01000, Digital Forensics Integrated Platform Development)

Acknowledgment

Grant: Development of digital forensics integrated platform

Supported by: Information and Communication Technology Promotion Center

References

1. B. Park, J. Park, and S. Lee, "Data concealment and detection in Microsoft Office 2007 files," Digital Investigation, vol. 5, no. 3-4, pp. 104-114, Mar. 2009. https://doi.org/10.1016/j.diin.2008.12.001
2. Z. Fu, X. Sun, Y. Liu, and B. Li, "Forensic investigation of OOXML format documents," Digital Investigation, vol. 8, no. 1, pp. 48-55, Jul. 2011. https://doi.org/10.1016/j.diin.2011.04.001
3. ECMA, "ECMA-376-1:2016 Office Open XML file format - fondamentals and markup language reference," ECMA International Publication, Oct. 2016.
4. E. Didriksen, "Forensic analysis of OOXML documents," MS. Thesis, Gjovik University College, 2014.
5. H. Langweg, "OOXML file analysis of the July 22nd terrorist manual," 13th International Conference on Communications and Multimedia Security, Sep. 2012.
6. SL Garfinkel and JJ Migletz, "New XML-based files implications for forensics," IEEE Security and Privacy, vol. 7, no. 2, Mar-Apr. 2009.

7. H. Chung, J. Park, and S. Lee, "Forensic analysis of residual information in adobe PDF files," Communications in Computer and Information Science, vol. 185, 2011.

8. YM Lee and S. Lee, "A Study for Forensic Methods of MS Excel Files," MS. Thesis, Korea University, 2015.

9. D. Jeong and S. Lee, "Study on the tracking revision history of MS Word files for forensic investigation," Digital Investigation, vol. 23, pp. 3-10, Dec. 2017. https://doi.org/10.1016/j.diin.2017.08.003