

사무용 SW 프로세스 메모리 데이터 분석

오 수 빈*, 손 태 식**
아주대학교 사이버보안학과 (학부생)*, (교수)**

Office software process memory data analysis

Subin Oh*, Taeshik Shon**
Dept. of Cybersecurity, Ajou University (Undergraduate Student)*, (Professor)**

요 약

코로나의 영향으로 업무환경, 수업환경 등이 비대면으로 전환된지 꽤 오랜 기간이 지났다. 이로 인해 사무용 SW에 대한 수요가 커지고, 업무와 학업을 불문하고 사무용 소프트웨어가 다양하게 사용되면서 우리 삶에 없어서는 안될 소프트웨어 중 하나가 되었다. 사무용 소프트웨어로 문서 작업을 하다보면, 여러 이유로 인해 입력 데이터가 소실되어 중요한 정보를 잃을 수가 있다. 백업 파일이 존재하지 않아도, 메모리에서 얻은 아티팩트를 통해 일부라도 복구할 수 있다면 매우 유용할 수 있다. 따라서 본 논문에서는 백업 파일이 존재하지 않은 경우를 가정하고 사무용 소프트웨어에서 작업했던 데이터를 프로세스 메모리 관점에서 분석하고자 한다. 연구를 위해 국내 점유율이 높은 Microsoft Office와 한컴오피스 제품군을 선택하여 Windows 10, 11 운영체제에서 분석하였다. 소프트웨어에 영어, 한글, 파일 암호 등의 데이터를 입력하고 파일을 저장하지 않아도 데이터가 프로세스 메모리에서 검출되는지 확인하는 방식으로 진행했으며, 마지막에는 운영체제에 따른 분석 결과의 차이를 비교하였다. 운영체제에 따른 차이는 없었으며 소프트웨어별로 프로세스 메모리에서 검출되는 입력 데이터의 종류에는 차이가 존재했다.

주제어 : 프로세스 메모리, 윈도우 10, 윈도우 11, Microsoft Office, 한컴오피스

ABSTRACT

It has been a long time since the work environment and class environment were switched to non-face-to-face due to the influence of COVID-19, which has become one of the indispensable software in our lives as demand for office software has increased and office software has been used in various fields. When working on documents with office software, input data may be lost for various reasons and important information may be lost. Even if a backup file does not exist, it can be very useful if it can be recovered in part through artifacts obtained from memory. Therefore, in this paper, assuming that a backup file does not exist, the data worked in the office software is analyzed from the perspective of process memory. For the study, Microsoft Office and Hancom Office product lines with high domestic market share were selected and analyzed in Windows 10 and 11 operating systems. It was conducted by entering data such as English, Korean, and file passwords into the software and checking whether data was detected in the process memory without storing files, and finally, the difference in analysis results according to the operating system was compared. There was no difference according to the operating system, and there was a difference in the type of input data detected in process memory by software.

Key Words : Process Memory, WINDOWS 10, WINDOWS 11, Microsoft Offices, Hancom Office

• Received 04 February 2022, Revised 04 February 2022, Accepted 31 March 2022
• 제1저자(First Author) : Subin Oh (Email : osb0408@ajou.ac.kr)
• 교신저자(Corresponding Author) : Taeshik Shon (Email : tsshon@ajou.ac.kr)

I. 서 론

COVID-19의 영향으로 대부분의 기업과 교육기관이 비대면 활동으로 전환하게 되면서 사무용 SW의 사용률이 급증했다. 일상 속에서 사무용 SW를 통해 문서 작업을 하다보면, 갑자기 전원이 끊어지거나 프로그램이 비정상적으로 종료되는 등의 이유로 의도치 않게 저장되지 않은 데이터를 잃는 경우가 있다. 사본이 있거나 백업 파일이 있는 경우에는 상관이 없지만, 그렇지 않은 경우에는 중요한 데이터를 잃게 된다. 따라서 이러한 상황을 예방할 방안에 대해 연구할 필요가 있고, 이를 프로세스 메모리 관점으로 접근해보고자 한다.

프로그램을 실행시키면, 프로세스는 종료되기 전까지 휘발성 메모리인 RAM에서 동작하게 된다. 프로세스가 종료되면 사용하던 자원을 반환하고 메모리에서 사라지므로 저장되지 않은 입력 데이터는 프로세스가 종료된 이후에 메모리에서 확인할 수 없다. 따라서 사무용 SW가 의도치 않게 종료되는 경우에 백업 파일이 없다면 파일이 저장되기 전에 데이터는 복구할 방법이 없다. 하지만 프로세스가 동작 중일 때 확인할 수 있는 프로세스 메모리에서는 저장되지 않은 입력 데이터도 확인할 수 있다는 특징이 있기 때문에 사무용 SW의 프로세스 메모리에서 확인할 수 있는 입력 데이터를 분석해볼 필요가 있다.

사무용 SW의 프로세스 메모리 데이터를 분석하기 위해서 국내 기준으로 사용률이 높은 제품군을 선정하여 실험을 진행했다. 2020년 기준으로 Microsoft Office의 국내 점유율은 약 70%이고 한컴오피스의 국내 점유율이 나머지 30%를 가진다[1]. 새로운 사무용 SW가 나오고는 있지만, 국내 점유율은 Microsoft Office와 한컴오피스 제품이 대부분을 차지한다. 본 연구에서는 Windows 10, Windows 11 운영체제에서 한글, 영어, 파일암호 등 다양한 입력 데이터가 프로세스 메모리에서 발견되는지에 대한 분석을 진행하고, 소프트웨어와 운영체제 사이에서 발생하는 차이를 도출한다.

본 논문의 구성은 다음과 같다. 2장에서 관련 연구 및 배경 지식에 대해서 살펴본다. 3장에서는 분석 환경 및 시나리오에 대해서 서술한다. 4장에서는 사무용 소프트웨어들을 중심으로 프로세스 메모리 데이터 분석을 진행하고 추가적으로 프로세스를 종료했을 때, 프로세스 메모리에서 확인한 데이터가 전체 메모리에서도 검출되는지 확인한다. 5장에서는 전체 결과를 바탕으로 각 소프트웨어, 운영체제마다 어떠한 차이가 있는지 설명한 뒤, 마지막으로 6장에서는 결론 및 향후 연구로 마무리한다.

II. 관련 연구 및 배경 지식

2.1 관련 연구

메모리에 대한 연구는 다양한 관점에서 진행되었다. 먼저 메모리를 수집하거나 분석하는 방법에 대한 연구들이 진행되었다. Sunu Thomas et al. 연구팀은 Windows 7 운영체제를 중심으로 메모리에서 아티팩트를 수집하는 연구를 진행했다. KPCR, PEB 등과 같은 Windows 커널 데이터 구조에 대해서 논의하고, Windows 7 32/64-bit의 메모리 정보에서 실행 중인 프로세스들의 목록을 출력하는 방법론을 제시하였다. 실행됐던 프로세스 목록은 추출할 수 있었지만 해당 프로세스들의 세부 데이터는 얻을 수 없었다는 한계가 있다[2]. Nicolas Ruff et al. 연구팀은 무료로 사용 가능한 도구들을 활용하여 live 메모리를 수집하는 방법에 대해 제시하였다. 무료로 사용 가능한 live 메모리 수집 도구와 분석 도구를 실제 사례와 함께 설명하였다. 또한 안티포렌식 방법들을 소개하였다[3]. Ralpa Palutke et al. 연구팀은 보안 분석가 관점에서 메모리에 접근할 수 없도록 하는 새로운 세 가지 방법을 제시한다. Windows와 Linux 환경에서 연구를 진행했으면 스스로의 기술들을 탐지하기 위한 몇 가지 접근법을 논의하기도 하였다[4].

또한 메모리에서 수사의 정황 증거 관점으로 아티팩트를 수집하는 연구들도 진행되었다. Jung-heum Park et al. 연구팀은 Microsoft Office의 PowerPoint를 중심으로 디지털 포렌식 관점에서 의미 있는 도출하는 연구를 진행했다. 파일 형식, 잉여 정보, 개체 식별자 등을 조사하는 방법을 제시했다. 또한 해당 연구에서 알아낸 포렌식 방식을 활용하여 가상의 사건을 통해 사건의 정황증거로 활용될 수 있음을 보였다[5]. S. Park and S. Lee는 HWP 파일을 중점으로 편집 이력과 편집 S/W 식별 등의 아티팩트를 도출하는 연구를 진행했다. 해당 연구에서는 문서 편집 S/W 식별, 문서 편집에 참여한 사용자에 대한 정보 등의 아티팩트를 찾아냈을뿐만 아니라, 문서 개체에 존재하는 InstanceId가 현재 시간에 기반한 값을 사용하고, 이를 시간 변환 알고리즘을 통해 개체 생성 시간을 역으로 계산하면 편집 이력을 확인하는 용도로 사용 가능함을 최초로 소개하였다. 또한 실제로 HWP 파일에 포렌식 아티팩트를 이용하여 수사를 진행할 사례를 소개하였다[6]. Yeonjae Im et al. 연구팀은 전자문서가 사건의 단서가 될 수 있다는 점을 고려하여 Microsoft Office 프로그램에서 작업 이력을 분석하는 연구를 진행했다. Microsoft Office 제품군 중 Word, Excel,

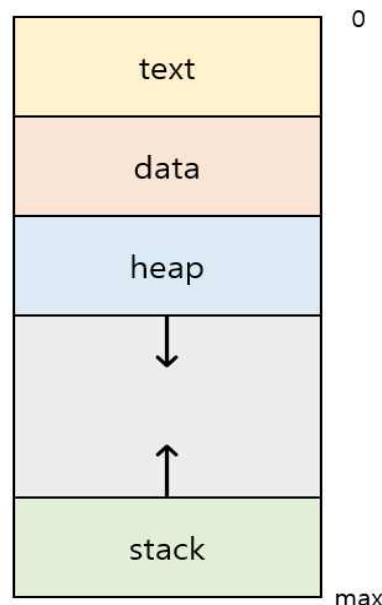
PowerPoint를 대상으로 이벤트 로그를 분석 결과를 도출했다. 또한 연구 결과를 통해 도출된 이벤트 로그들을 가상의 시나리오를 설정하여 활용되는 사례를 제시하였다[7]. 이외에도 메모리 맵 파일 관점으로 덤프된 메모리에서 파일을 카빙하는 방법, Linux에서 프로세스 사용자 공간에 포함되는 데이터 분석, 프로그램의 일부를 덮어쓰으로써 프로세스의 메모리 크기를 줄이는 기술 등에 대한 연구가 진행되었다[8-10].

메모리 데이터 관점의 접근은 아니지만, 사무용 SW에서 얻을 수 있는 아티팩트를 분석한 선행 연구가 존재한다. YeonJoo Lee et al. 연구팀은 Polaris Office에서 문서 열람, 문서 편집 등 특정 행위를 실행하는 시나리오를 설정하여 실행한 뒤 발견할 수 있는 아티팩트에 대해 연구한 바가 있다. Windows 10과 macOS 두 운영체제에서 최근 사용된 파일 목록, 사용자 계정 정보 등 과 같은 다양한 아티팩트를 도출하였으며 운영체제에 따른 아티팩트의 차이를 결과로 도출하였다[11].

본 연구에서는 사무용 SW를 프로세스 메모리로 방식으로 접근하여, 프로세스 메모리 상에 입력 데이터 보존 정도를 분석한다는 차별점이 존재한다. 또한 선행 연구들은 공통적으로 Windows 11 운영체제에 대한 연구 결과는 없다는 한계가 존재하기 때문에, Windows 11 운영체제에서도 사무용 SW에 관련된 연구를 진행하여 Windows 10과 다른 점을 확인해볼 필요가 있다.

2.2 배경 지식

RAM(Random Access Memory)는 휘발성 메모리로, PC의 전원이 차단되면 데이터가 사라진다는 특징이 있다. RAM에는 브라우저 기록, 사용한 command, 로드된 드라이버, 실행중인 프로세스 등에 대한 정보가 담겨있다[9]. 따라서 메모리 데이터를 분석하면 악성 코드나 랜섬웨어에 감염되었을 경우, 사건 전후 메모리를 비교함으로써 문제를 해결할 수 있다. 또한 메모리에 남겨진 아티팩트들이 수사에 정황 증거로도 사용될 수 있다. 프로세스(Process)는 메모리에서 실행 중인 프로그램을 의미한다. 프로세스의 구조는 [그림 1]과 같이 text, data, heap, stack으로 구성되어 있다. 프로세스는 RAM에서 동작하며, 프로세스가 종료되면 사용하던 자원을 반납하고 종료된다. 그렇기 때문에 프로세스가 종료된 이후에는 저장되지 않은 입력 데이터는 메모리에서 확인할 수 없다.



〈Figure 1〉 Process Memory Structure

따라서 본 논문에서는 프로세스가 종료되기 전의 데이터를 분석할 필요가 있으므로 프로세스 메모리 덤프를 진행한다. Windows 운영체제는 작업 관리자에서 프로세스 메모리 덤프 기능을 제공함으로 해당 기능을 이용하여 프로세스 메모리 덤프를 수행한다. 작업 관리자 기능을 통해 프로세스 메모리 덤프를 수행하면, 'C:\Users\Username\AppData\Local\Temp' 경로에 '프로세스명.DMP' 파일로 저장된다. 연구의 분석 대상인 사무용 SW들의 프로세스 이름은 [표 1]과 같다. 프로세스명은 Windows 10과 Windows 11에서 동일하게 확인된다.

〈Table 1〉 Process Name

Product	Program	Process Name	Version
Microsoft Office	Word	WINWORD	2202
	Excel	EXCEL	
	PowerPoint	POWERPNT	
한컴오피스 2020	한워드	Hword	11.0.0.5125
	한글	Hwp	11.0.0.5178
	한쇼	HShow	11.0.0.5142
	한셀	HCell	11.0.0.5334

III. 분석 환경 및 시나리오

본 논문에서는 [표 2]와 같은 환경에서 연구를 진행하였다. Windows 10 Home과 Windows 11 Pro 운영체제에서 진행하였으며, Windows 11 Pro는 가상환경을 이용했다. 프로세스 메모리 덤프는 작업 관리자의 기능을 사용했으며, 덤프 파일은 HxD를 사용하여 분석하였다.

〈Table 2〉 Analysis environment

Type	OS/Tool	Version
OS	Windows 10 Home	19042
	Windows 11 Pro(virtual machine)	21996.1
Analysis tool	HxD	2.5.0.0

분석 대상 소프트웨어와 시나리오에 사용된 입력 데이터는 [표 3]과 같다. 분석 대상은 국내 점유율이 높은 Microsoft Office와 한컴오피스 제품군을 선정했다. 저장되지 않은 문서 데이터의 복구 가능성을 확인하기 위하여 모든 입력 데이터들은 프로그램의 백업 기능을 끈 채 저장하지 않고 입력되었다. 입력 데이터 유형은 영어, 한글, 표 내부 영어, 파일 암호 설정, 이미지 파일로 선정했으며, 엑셀의 경우에는 수식도 추가로 입력했다. 입력 데이터는 [표 4]와 같이 설정했다. 입력 데이터 유형의 경우에는 입력 빈도가 높다고 판단되는 유형들을 선정했다.

〈Table 3〉 Target office software

Product	Program Name	Input Data	Version
Microsoft Office	Word	영어/한글 텍스트, 표 내부 영어 텍스트, 파일 암호, 이미지	18.2106.12410.0
	PowerPoint		
	Excel	영어/한글 텍스트, 수식, 파일 암호, 이미지	
한컴오피스 2020	한워드	영어/한글 텍스트, 표 내부 영어 텍스트, 파일 암호, 이미지	11.0.0.5125
	한글		11.0.0.5178
	한쇼		11.0.0.5142
	한셀	영어/한글 텍스트, 수식, 파일 암호, 이미지	11.0.0.5334

〈Table 4〉 Input data

Type	English	Korean	English in table	File password	Image	Formular
Data	Mem0ry	메모리 포렌식	F0reN2ic	f00rensic	tuLip2.jpg	=RAND()*10

IV. 프로세스 메모리 데이터 분석

4.1. Windows 10 Home

4.1.1. Microsoft Office

a. Word

[그림 2]부터 [그림 6]까지는 Windows 10 Home 버전에서 Microsoft Office의 Word에서 프로세스 메모리를 분석한 결과이다. 분석 결과, [그림 3]에서 영어 입력 데이터인 'Mem0ry'를 확인할 수 있었다. [그림 3]은 한글 입력 데이터인 '메모리 포렌식'을 확인할 수 있었다. 메모리 포렌식은 utf-8로 인코딩하면 'EB A9 94 EB AA A8 EB A6 AC 20 ED 8F AC EB A0 8C EC 8B 9D'이 된다. [그림 4]는 표 내부에 입력했던 'F0reN2ic'를 확인할 수 있었다. [그림 5] 문서의 암호를 설정할 때 입력했던 'f00rensic'이 프로세스 메모리에 암호화 없이 그대로 노출되어있는 것을 확인할 수 있었다. [그림 6]은 첨부했던 이미지 파일의 파일명을 확인할 수 있었다. 프로세스 메모리의 정보를 통해 첨부되었던 이미지의 파일을 추출할 수는 없었다.

```
00 00 00 51 22 00 08 8A FB DD 66 06 00 00 00 4D ...Q"..ŠŭŸf....M
65 6D 30 72 79 00 00 91 5B 00 08 0A 00 00 00 F9 em0ry..`[.....ù
```

〈Figure 2〉 English data in word with windows 10

```
7B 5C 22 74 65 78 74 5C 22 3A 5C 22 EB A9 94 EB {"text\":"\ë@"è
AA A8 EB A6 AC 20 ED 8F AC EB A0 8C EC 8B 9D 5C "a"è|~ í.~ë (i<.\
5C 72 5C 22 2C 5C 22 74 65 78 74 55 6E 69 74 5C \r\","textUnit\
```

〈Figure 3〉 Korean data in word with windows 10

```
22 3A 5B 7B 5C 22 74 65 78 74 5C 22 3A 5C 22 46 ":[{"text\":"\F
30 72 65 4E 32 69 63 20 5C 22 2C 5C 22 74 65 78 0reN2ic \","tex
```

〈Figure 4〉 In-table data in word with windows 10

```
FC A5 8E BA 01 00 00 33 85 86 64 FE 7C 01 8C 66 ůŽž°...3...tdp|.Ef
00 30 00 30 00 72 00 65 00 6E 00 73 00 69 00 63 .0.0.r.e.n.s.i.c
```

〈Figure 5〉 Password data in Excel with Windows 10

```
00 5C 00 74 00 75 00 4C 00 69 00 70 00 32 00 2E .\..t.u.L.i.p.2..
00 6A 00 70 00 67 00 00 00 62 40 06 00 01 07 06 .j.p.g...b@.....
```

〈Figure 6〉 Image file name in word with windows 10

b. Excel

[그림 7]부터 [그림 10]까지는 Windows 10 Home 버전에서 Microsoft Office의 Excel에서 프로세스 메모리를 분석한 결과이다. Excel에서는 영어, 파일 암호, 이미지 파일명, 수식을 확인할 수 있었다.

```
00 00 00 00 06 00 00 00 06 00 4D 00 65 00 6D 00 .....M.e.m.
30 00 72 00 79 00 01 00 00 00 00 00 00 00 06 00 0.r.y.....
```

〈Figure 7〉 English data in Excel with windows 10

```

0C 94 D4 3C 00 6A 00 8C 09 00 66 00 30 00 30 00 . "Ô<.j.Æ..f.0.0.
72 00 65 00 6E 00 73 00 69 00 63 00 00 00 00 00 r.e.n.s.i.c.....

```

〈Figure 8〉 Password data in Excel with windows 10

```

E2 2F 69 00 74 00 75 00 4C 00 69 00 70 00 32 00 â/i.t.u.L.i.p.2.
2E 00 6A 00 70 00 67 00 00 00 1A 00 00 00 00 00 ..j.p.g.....

```

〈Figure 9〉 Image file name in word with windows 10

```

3E D4 65 3F 75 06 00 90 0A 00 3D 00 52 00 41 00 >Ôe?u.....=.R.A.
4E 00 44 00 28 00 29 00 2A 00 31 00 30 00 00 00 N.D.(.)*.1.0...

```

〈Figure 10〉 Formula data in Excel with windows 10

c. PowerPoint

[그림 11]부터 [그림 14]까지는 Windows 10 Home 버전에서 Microsoft Office의 PowerPoint에서 프로세스 메모리를 분석한 결과이다. PowerPoint에서는 영어, 한글, 표 내부 데이터, 이미지 파일명을 확인할 수 있었다. PowerPoint는 같은 제품군인 Word, Excel과 달리 암호를 확인할 수 없었다.

```

A1 69 CF 00 80 4D 00 65 00 6D 00 30 00 72 00 79 ;iï.€M.e.m.0.r.y

```

〈Figure 11〉 English data in PowerPoint with windows 10

```

22 2F 3E 3C 61 3A 74 3E EB A9 94 EB AA A8 EB A6 "/><a:t>ë@"ë"ë!
AC 20 ED 8F AC EB A0 8C EC 8B 9D 3C 2F 61 3A 74 1.~ë Eik.</a:t

```

〈Figure 12〉 Korean data in Powerpoint with windows 10

```

22 3A 5C 5C 5C 22 46 30 72 65 4E 32 69 63 5C 5C ":\\\\"F0reN2ic\\

```

〈Figure 13〉 In-table text in PowerPoint with windows 10

```

70 00 5C 00 74 00 75 00 4C 00 69 00 70 00 32 00 p.\.t.u.L.i.p.2.
2E 00 6A 00 70 00 67 00 00 00 69 00 70 00 32 00 ..j.p.g...i.p.2.

```

〈Figure 14〉 Image file name in word with windows 10

4.1.2. 한컴오피스 2020

a. 한글

[그림 15]부터 [그림 17]까지는 Windows 10 Home 버전에서 한컴오피스 2020의 한글에서 프로세스 메모리를 분석한 결과이다. 한글에서는 영어, 표 내부 데이터, 파일 암호, 이미지 파일명을 확인할 수 있었다.

```

00 00 00 06 00 00 00 07 00 00 00 4D 00 65 00 6D .....M.e.m
00 30 00 72 00 79 00 00 00 00 00 00 00 00 50 .0.r.y.....P

```

〈Figure 15〉 English text data in hwp with windows 10


```

00 00 00 00 00 00 00 00 00 00 00 46 00 30 00 72 .....F.0.r
00 65 00 4E 00 32 00 69 00 63 00 0F 00 00 00 0E .e.N.2.i.c.....

```

〈Figure 16〉 In-table text in hwp with windows 10

```

00 00 00 66 00 30 00 30 00 72 00 65 00 6E 00 73 ...f.0.0.r.e.n.s
00 69 00 63 00 00 00 01 00 00 00 00 00 00 02 .i.c.....

```

〈Figure 17〉 Password data in hwp with windows 10

```

84 B9 3A 00 20 00 74 00 75 00 4C 00 69 00 70 00 „¹:. .t.u.L.i.p.
32 00 2E 00 6A 00 70 00 67 00 0D 00 0A 00 D0 C6 2...j.p.g.....ĐÆ

```

〈Figure 18〉 Image file name in word with windows 10

b. 한셀

[그림 19]부터 [그림 22]까지는 Windows 10 Home 버전에서 한컴오피스 2020의 한셀에서 프로세스 메모리를 분석한 결과이다. 한셀에서는 영어, 파일 암호, 이미지 파일명, 수식을 확인할 수 있었다.

```

00 48 1E 84 41 75 10 00 88 22 00 4D 00 65 00 6D .H.„Au..^\".M.e.m
00 30 00 72 00 79 00 22 00 20 00 85 C7 25 B8 28 .0.r.y.\". ....Ç%, (

```

〈Figure 19〉 English text data in HanCell with windows 10

```

88 66 00 30 00 30 00 72 00 65 00 6E 00 73 00 69 ^f.0.0.r.e.n.s.i
00 63 00 00 00 00 00 00 00 00 00 00 00 00 00 .c.....

```

〈Figure 20〉 Password data in HanCell with windows 10

```

00 00 00 00 74 00 75 00 4C 00 69 00 70 00 32 00 ....t.u.L.i.p.2.
2E 00 6A 00 70 00 67 00 00 00 00 00 00 00 00 .j.p.g.....

```

〈Figure 21〉 Image file name in word with windows 10

```

88 3D 00 52 00 41 00 4E 00 44 00 28 00 29 00 2A ^=.R.A.N.D.(.) .*
00 31 00 30 00 00 00 00 00 00 00 00 00 00 00 .1.0.....

```

〈Figure 22〉 Formular data in HanCell with windows 10

c. 한쇼

[그림 23]부터 [그림 26]까지는 Windows 10 Home 버전에서 한컴오피스 2020의 한쇼에서 프로세스 메모리를 분석한 결과이다. 한쇼에서는 영어, 표 내부 데이터, 파일 암호, 이미지 파일명을 확인할 수 있었다.

```

00 00 00 00 00 00 00 00 E4 24 A5 71 0E 00 00 00 4D .....ä$¥q....M
00 65 00 6D 00 30 00 72 00 79 00 20 00 54 BA A8 .e.m.0.r.y. .T°

```

〈Figure 23〉 English data in HanShow with windows 10

```

00 00 00 46 00 30 00 72 00 65 00 4E 00 32 00 69 ...F.0.r.e.N.2.i
00 63 00 00 00 00 00 00 00 00 00 00 00 00 00 .c.....

```

〈Figure 24〉 In-table data in HanShow with windows 10

```

00 00 00 0F 00 00 00 66 00 30 00 30 00 72 00 65 .....f.0.0.r.e
00 6E 00 73 00 69 00 63 00 00 00 01 00 00 00 00 .n.s.i.c.....

```

〈Figure 25〉 Password data in HanShow with windows 10

```

00 6B 00 74 00 6F 00 70 00 5C 00 74 00 75 00 4C .k.t.o.p.\.t.u.L
00 69 00 70 00 32 00 2E 00 6A 00 70 00 67 00 00 .i.p.2...j.p.g..

```

〈Figure 26〉 Image file name in word with windows 10

d. 한워드

[그림 27]부터 [그림 30]까지는 Windows 10 Home 버전에서 한컴오피스 2020의 한워드에서 프로세스 메모리를 분석한 결과이다. 한워드에서는 영어, 표 내부 데이터, 파일 암호, 이미지 파일명을 확인할 수 있었다.

```

24 04 88 4D 00 65 00 6D 00 30 00 72 00 79 00 0A $.~M.e.m.o.r.y..

```

〈Figure 27〉 English data in HanWord with windows 10

```

00 08 D7 0C 1A 65 74 E7 E8 00 27 04 8A 46 00 30 ..x...etçè.'.šF.0
00 72 00 65 00 4E 00 32 00 69 00 63 00 0A 00 44 .r.e.N.2.i.c...D

```

〈Figure 28〉 In-table data in HanShow with windows 10

```

00 00 00 E4 24 A5 71 09 00 00 00 66 00 30 00 30 ...ä$¥q....f.0.0
00 72 00 65 00 6E 00 73 00 69 00 63 00 00 00 00 .r.e.n.s.i.c....

```

〈Figure 29〉 Password data in HanWord with windows 10

```

00 6B 00 74 00 6F 00 70 00 5C 00 74 00 75 00 4C .k.t.o.p.\.t.u.L
00 69 00 70 00 32 00 2E 00 6A 00 70 00 67 00 00 .i.p.2...j.p.g..

```

〈Figure 30〉 Image file name in word with windows 10

4.2. Windows 11 Pro

Windows 11에서 Windows 10과 동일한 방식으로 분석해본 결과, 제품군에 상관없이 Windows 10에서 분석한 결과와 동일하게 나왔다.

4.3. 전체 메모리와 프로세스 메모리 비교

모든 경우에서 프로세스 메모리에서 확인되었던 데이터가, 프로세스에 저장되지 않고 종료된 이후 전체 메모리에서 똑같이 검출되는지 확인해봤다. 그 결과 모든 데이터 유형이 전체 메모리에서는 확인되지 않았다. 따로 저장되거나 백업본이 존재하지 않기 때문에 전체 메모리에서는 확인할 수 없는 것으로 추측된다.

V. 결과 및 토의

본 논문에서는 사무용 SW에서 저장되지 않은 입력 데이터가 프로세스 메모리에 존재하는지 확인했다. 분석 대상으로 Microsoft Office의 Word, Excel, PowerPoint와 한컴오피스의 한글, 한셀, 한쇼, 한워드를 선정했다. 목표는 Windows 10, 11 운영체제에서 백업 설정을 끄고 파일을 저장하지 않은 상태에서 입력 데이터 유형별로 데이터가 프로세스 메모리에 남는지 확인하는 것이다. 실험 결과는 [표 5]와 같으며, 운영체제에 상관없이 동일한 결과를 얻었다. 그러나 소프트웨어마다 확인되는 데이터 유형은 다양했다. 영어 데이터와 이

미지 파일명의 경우에는 모든 소프트웨어에서 확인되었으며, 파일 암호의 경우에는 Microsoft Office 제품군의 PowerPoint를 제외하고는 모두 암호화 되지 않은 상태로 메모리에 그대로 노출되었다. 또한 이미지 파일의 경우에는 첨부되었던 파일명을 제외하고는 특별한 아티팩트는 발견하지 못했다.

〈Table 5〉 The result of an analysis

Product	Program	Detected Data
Microsoft Office	Word	영어, 한글, 표, 암호, 이미지
	Excel	영어, 암호, 이미지, 수식
	PowerPoint	영어, 한글, 표, 이미지
한컴오피스 2020	한글	영어, 표, 암호, 이미지
	한셀	영어, 암호, 이미지, 수식
	한쇼	영어, 표, 암호, 이미지
	한워드	영어, 표, 암호, 이미지

자주 사용되는 입력 데이터 유형 위주로 저장된 데이터가 아니라도, 프로세스 메모리에서 검출됨을 확인했다. 이 데이터들은 프로세스가 종료된 이후에 전체 메모리에서는 확인되지 않고, 프로세스 메모리에서만 확인했기 때문에 사무용 SW에서 프로세스 메모리에 보다 가치 있는 데이터가 담겨있다고 할 수 있다.

VI. 결론

코로나의 영향으로 사무용 SW의 사용률은 높아지고 있다. 그렇기 때문에 백업 파일 없이 데이터가 소실될 경우, 프로세스 메모리 데이터만 획득할 수 있다면 프로그램 특성에 맞게 데이터를 일부 복원할 수 있다. 프로세스 메모리 데이터를 수집할 방법을 모색하기 전에, 프로세스 메모리에 가치 있는 데이터가 있음을 확인할 필요가 있었다. 4장에서 Windows 10과 Windows 11 환경에서 프로그램마다 입력 데이터 유형별로 프로세스 메모리에 데이터가 남게 되는지를 분석하고 전체 메모리와 프로세스 메모리에서 검출되는 데이터도 차이도 확인했다. 5장에서는 결과를 정리하여, 운영체제 사이에는 차이가 없었지만 프로그램마다 발견되는 입력 데이터 유형이 다름을 확인했다. 종료된 프로세스의 경우, 전체 메모리에서는 확인되지 않은 데이터가 프로세스 메모리에서는 확인되었다.

본 연구에서는 한국어와 영어를 중심으로 텍스트 데이터를 분석하였지만, 추후에는 다른 언어들과 함께 분석 데이터 종류, 분석 프로그램 종류를 더 다양화한 연구도 수행해볼 것이다. 또한 사무용 SW가 갑작스럽게 종료되었지만 백업 파일이 없는 경우 프로세스 메모리를 활용해볼 가치가 있음을 확인하였기 때문에 프로세스가 종료된 이후에도, 프로세스 메모리 데이터를 수집할 방법을 모색할 것이다.

참 고 문 헌 (References)

- [1] DONG-A ILBO, "Polaris Office Cloud Version Challenges Domestic Office SW Market". Available: <https://www.donga.com/news/It/article/all/20200414/100654865/1>. 2022.1.25. confirmed.
- [2] S. Thomas, K. K. Sherly and S. Dija, "Extraction of memory forensic artifacts from windows 7 RAM image," 2013 IEEE Conference on Information & Communication Technologies, 2013, pp. 937-942, doi: 10.1109/CICT.2013.6558230.
- [3] Ruff, N. (2008). Windows memory forensics. Journal in Computer Virology, 4(2), 83-100.
- [4] Ralph Paluctke, Frank Block, Patrick Reichenberger, Dominik Stripeika, Hiding Process Memory Via Anti-Forensic Techniques, Forensic Science International: Digital Investigation, Volume 33, Supplement, 2020, 301012, ISSN 2666-2817, doi: 10.1016/j.fsidi.2020.301012.
- [5] Jung-heum Park ·SungMoon Hong ·YongSeok Choi ·SangJin Lee, "Methods for investigating MS PowerPoint Files from the Viewpoint of Digital Forensics", Journal of digital forensics, 2(1), pp.13-28, Dec. 2008.
- [6] S. Park and S. LEE, "Forensic Investigation of HWP File," Journal of Digital Forensics, vol. 14, no. 4, pp. 408 - 425, Dec. 2020.
- [7] Yeonjae Im ·Jungheum Park ·Sangjin Lee, "Forensic Exploration of Microsoft Office's Diagnostic Logs", vol. 15, no 2, pp. 24 - 34, Dec. 2021.
- [8] Van Baar, R. B., Alink, W., & Van Ballegooij, A. R. (2008). Forensic memory analysis: Files mapped in memory. digital investigation, 5, S52-S57.
- [9] Block, F., & Dewald, A. (2017). Linux memory forensics: Dissecting the user space process heap. Digital Investigation, 22, S66-S75.
- [10] R. L. Bowman, E. J. Ratliff and D. B. Whalley, "Decreasing process memory requirements by overlapping program portions," Proceedings of the Thirty-First Hawaii International Conference on System Sciences, 1998, pp. 115-124 vol.7, doi: 10.1109/HICSS.1998.649204.
- [11] YeonJoo Lee ·JeongMin Kim ·SungJin Lee, "A Study of Polaris Office Forensic Artifact", Journal of Digital Forensics, vol. 14, no. 4, pp. 368 - 378, Dec. 2020.

저 자 소 개



오 수 빈 (Subin Oh)

준회원

2019년 3월 ~ 현재 : 아주대학교 사이버보안학과 학사과정

관심분야 : Cyber Security, Digital Forensics



손 태 식 (Taeshik Shon)

평생회원

2000년 : 아주대학교 정보및컴퓨터공학부 졸업(학사)

2002년 : 아주대학교 정보통신전문대학원 졸업(석사)

2005년 : 고려대학교 정보보호대학원 졸업(박사)

2004년 ~2005년 : University of Minnesota 방문연구원

2005년 ~2011년 : 삼성전자 통신· DMC 연구소 책임연구원

2017년 ~2018년 : Illinois Institute of Technology 방문교수

2011년 ~ 현재 : 아주대학교 정보통신대학 사이버보안학과 교수

관심분야 : Digital Forensics, ICS/Automotive Security