

악성 문서의 '간단한' 정적분석 가이드 라인

중부대 SCP 정보보안 동아리
92113566 1학년 부원 노무승

목차

0-0. 개요

0-1. 문서형 악성코드

1. OLE 파일 구조

2. Microsoft 악성 문서 진단
(PPT, XLS, Docx)

3. PDF 악성 문서 진단

4. 아래 한글 악성 문서 진단
(HWP 5.x / HWPML)

0-0. 개요

보안 프로그래밍 과목 저희 팀 팀플 주제가
<악성 문서 진단 프로그램>임.

따라서 실제 프로그램 구현 보다는
정적분석 가이드 라인 제시(기획)에 초점을 두었음.

0-1. 문서형 악성코드

일반적인 실행 파일이 아닌 문서 구조를 띤 악성코드. APT 공격에 주로 사용됨.

북한 관련 내용으로 위장한 악성 문서 유포

A 권영수 기자 | © 입력 2020.12.09 22:18 | @ 댓글 0

이스트시큐리티, 북한 연계 APT 조직 공격으로 추정
북한동향 PDF 문서파일 링크 클릭 시 이메일 계정 해킹 우려

[아이티데일리] 최근 통일부 자료와 통일 관련 공모전 신청서를 사칭한 악성 문서가 유포되고 있어 각별한 주의가 요구된다. 이스트시큐리티는 해당 공격을 '탈북'과 '금성121' 등 북한 연계 조직의 소행으로 추정하고 있다.

9일 이스트시큐리티(대표 정성환)는 북한 이슈를 악용한 APT 공격이 언어어 발생하고 있다며, 주의를 당부했다.

이스트시큐리티 시큐리티대응센터(이하 ESRC)가 이변에 발견한 APT 공격은 ▲통일부 사칭 악성 이메일 공격과 ▲평화 통일 관련 이가 공모전 신청서를 사칭한 악성 HWP 문서 공격이다.

통일부 사칭 이메일과 통일 관련 공모전 신청서 사칭한 HWP 파일(제출: 이스트시큐리티)

‘사례비 지급의뢰서’…북한, 목표 대상에 악성 MS워드 문서 발송

By 문동희 기자 - 2021.08.27 3:40 오후

본 악성서에는 명시된 거래은행, 계좌번호, 계좌번호에 의하여 대금
을 계좌 입금한 이후에 발생한 손해에 대하여 동국대는 일체 그
책임을 지지 않습니다.

【개인정보 수집 · 이용 동의서】
개인정보보호법 제33조(개인정보의 수집 · 이용, 제33조)에 따라 받는 정보에
따라 동의합니다. 귀하의 개인정보는 수집 목적 외 다른 목적으로는 이용
하지 않으며, 정보주체가 개인정보에 대한 열람, 정정, 삭제, 지각정지, 이의제기
하고자 할 때에는 개인정보보호책임자를 통해 요구할 수 있으며, 개인정보보호
시 개인정보처리방침에 명시된 권익침해 구제방법을 통해 구제받을 수 있습니
다.

또 귀하는 상기 개인정보 수집 · 이용에 대한 동의를 거부할 권리가 있으나 미동여
대금 지급에 차질이 있을 수 있습니다.

☐ 동의함 ☐ 동의하지 않음
20 년 월 일
약장동의자 (서명 또는 날인)
업무담당자 (서명 또는 날인)

북한 해킹 조직 '탈북'이 공격에 사용한 문서, 사용자가 콘텐츠 사용(붉은 네모) 버튼을 누르면 악성코드가 실행된다. / 사진=이스트시큐리티 시큐리티대응센터 제공

북한 해킹 조직 '탈북', PDF 취약점 공격 벌여

© 김선애기자 | © 승인 2021.08.03 10:29 | @ 댓글 0

본 악성서에는 명시된 거래은행, 계좌번호, 계좌번호에 의하여 대금
을 계좌 입금한 이후에 발생한 손해에 대하여 동국대는 일체 그
책임을 지지 않습니다.

【개인정보 수집 · 이용 동의서】
개인정보보호법 제33조(개인정보의 수집 · 이용, 제33조)에 따라 받는 정보에
따라 동의합니다. 귀하의 개인정보는 수집 목적 외 다른 목적으로는 이용
하지 않으며, 정보주체가 개인정보에 대한 열람, 정정, 삭제, 지각정지, 이의제기
하고자 할 때에는 개인정보보호책임자를 통해 요구할 수 있으며, 개인정보보호
시 개인정보처리방침에 명시된 권익침해 구제방법을 통해 구제받을 수 있습니
다.

또 귀하는 상기 개인정보 수집 · 이용에 대한 동의를 거부할 권리가 있으나 미동여
대금 지급에 차질이 있을 수 있습니다.

☐ 동의함 ☐ 동의하지 않음
20 년 월 일
약장동의자 (서명 또는 날인)
업무담당자 (서명 또는 날인)

북한 해킹 조직 '탈북'이 공격에 사용한 문서, 사용자가 콘텐츠 사용(붉은 네모) 버튼을 누르면 악성코드가 실행된다. / 사진=이스트시큐리티 시큐리티대응센터 제공

0-1. 문서형 악성코드

크게 다음과 같은 방식이 존재함.

- 문서 프로그램의 취약점을 이용
- 문서 프로그램에서 제공하는 스크립트 기능을 이용

본 가이드라인은 스크립트를 사용하는 악성코드를 대상으로 함.

1. OLE 파일 구조

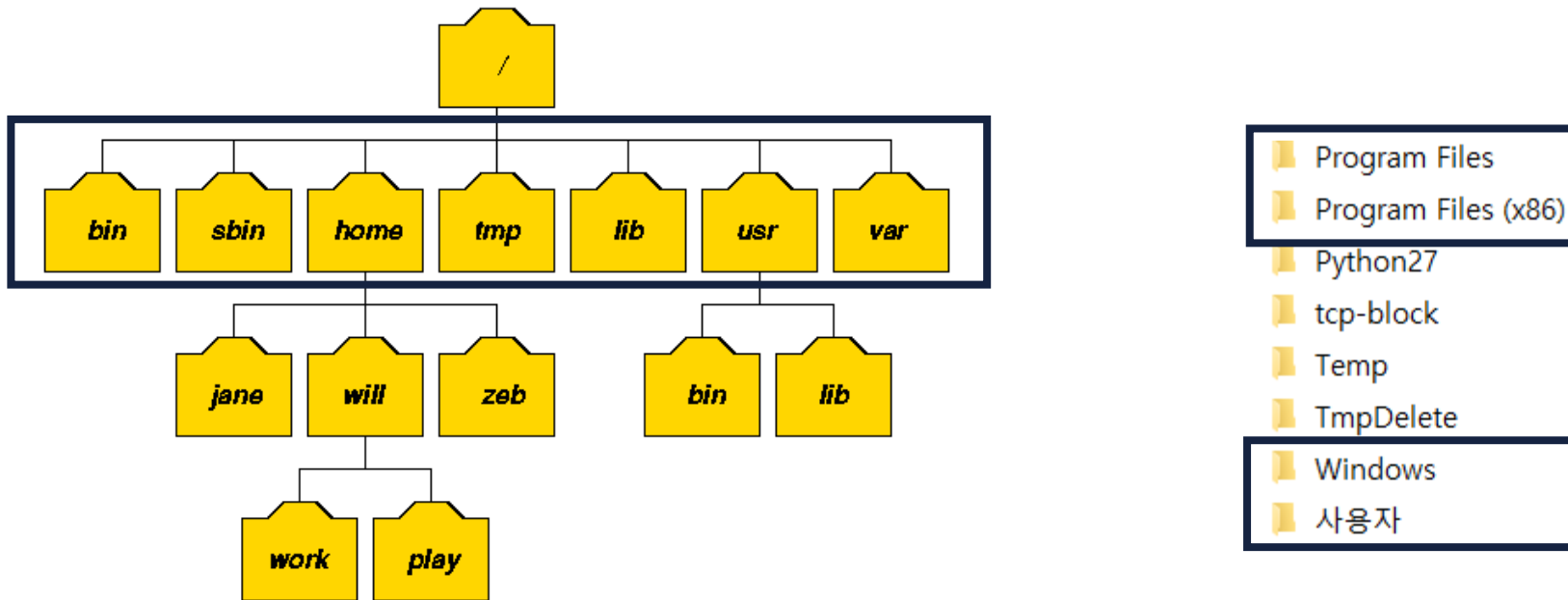
1-1. 파일 시스템 (FS)



파일 시스템				
마이크로소프트	FAT	NTFS	ReFS	
애플	HFS	HFS+	APFS	
리눅스/유닉스	ext2	ext3	ext4	
	Btrfs	F2FS	XFS	ZFS

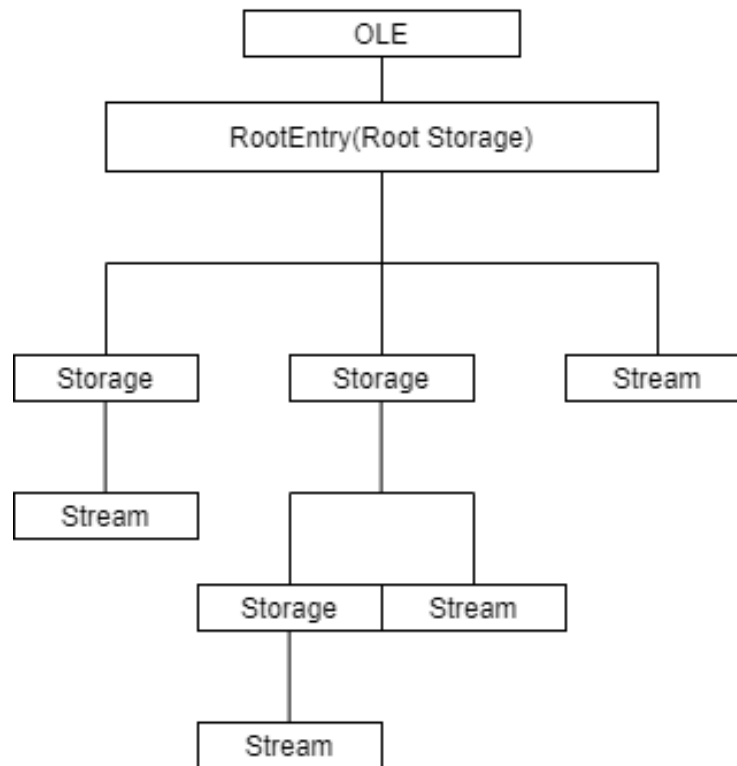
물리적인 디스크 내에서 데이터를 배치하고 관리하기 위한 체계

1-2. 루트 파일 시스템 (RFS)



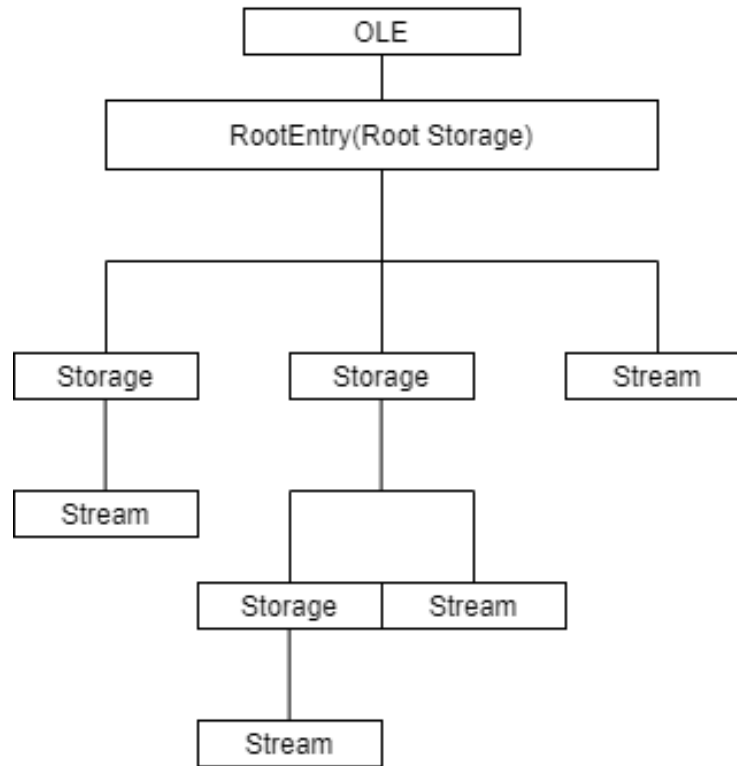
루트 디렉터리(/, C:₩)에서부터 데이터를 배치하고 관리하기 위한 체계

1-3. OLE 파일 구조



OLE(Object Linking and Embedding) 구조는
Microsoft에서 제작한 객체 연결 및 삽입이 용이한 구조임.

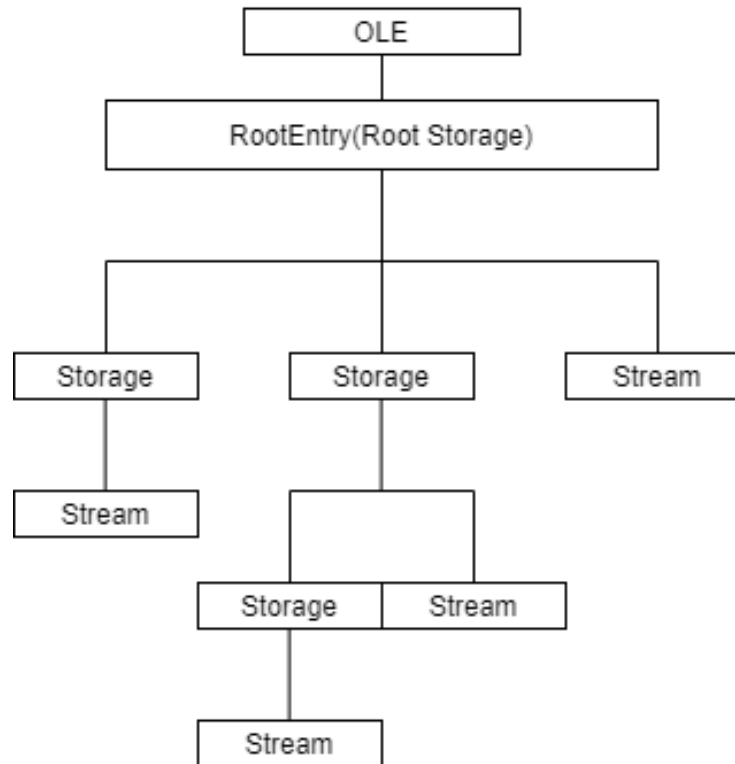
1-3. OLE 파일 구조



소규모의 파일 시스템과 같은 구조를 지니고 있어
뛰어난 호환성을 가지고 있다.

따라서 많은 파일들이 OLE 파일 형식을 사용하고 있다.

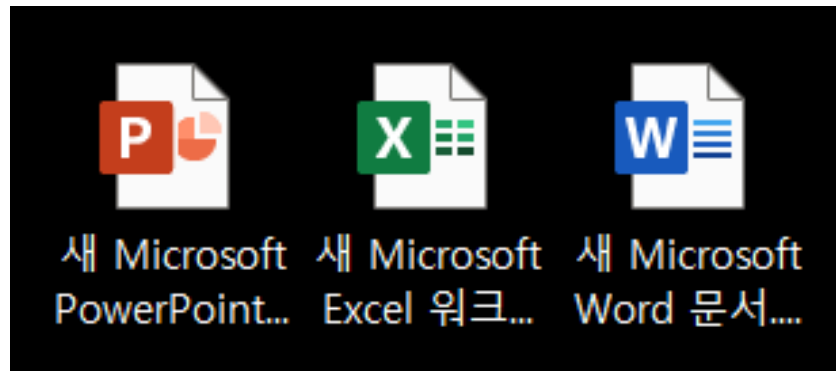
1-3. OLE 파일 구조



Storage는 폴더, Stream은 파일의 개념

2. Microsoft 악성 문서 진단

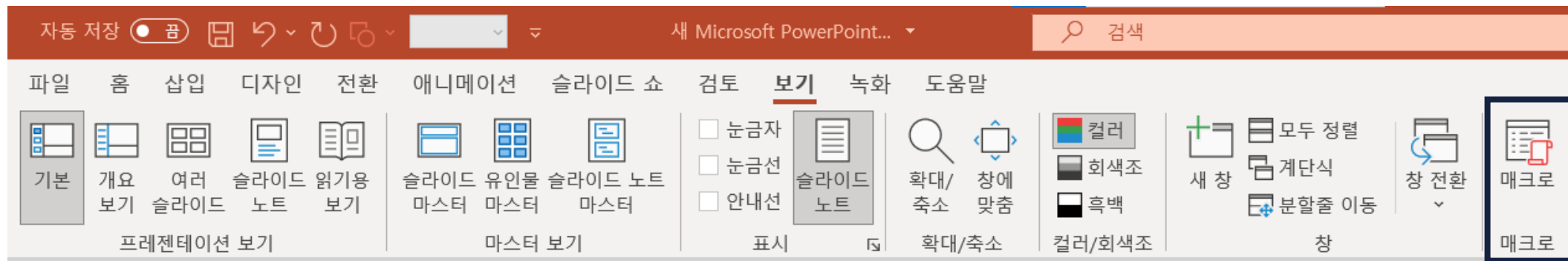
2-1. Microsoft 문서



크게 파워포인트(.ppt), 엑셀(.xlsx), 워드(.docx)로 분류할 수 있음.

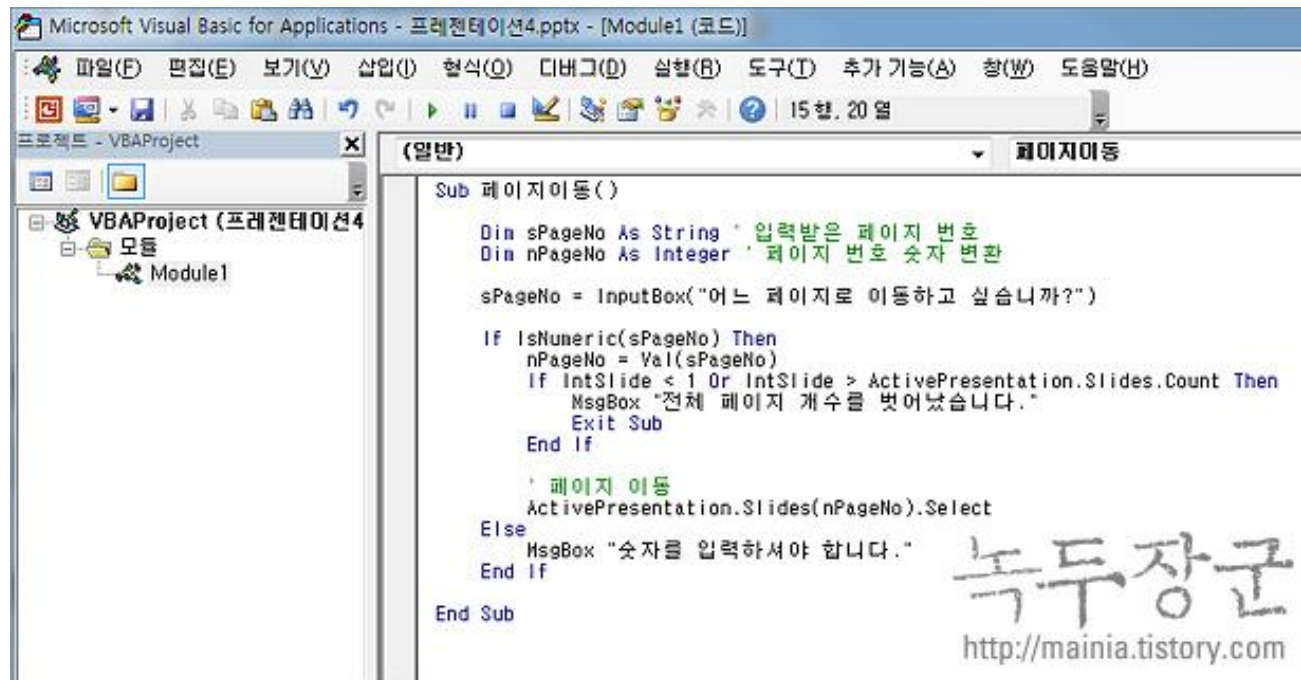
OLE 구조를 따름.

2-2. 어떻게 악성 문서가 되나



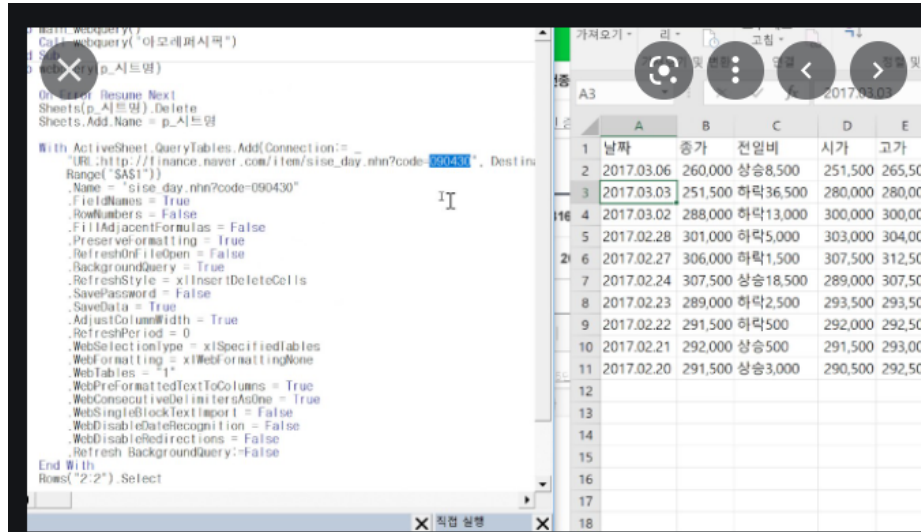
매크로 기능 때문.

2-3. 매크로 기능



VBA(Visual Basic Application) 프로그래밍 언어를 사용하며,
프로그래밍 언어를 이용해 내부 요소를 컨트롤 할 수 있음.

2-4. 매크로를 왜 사용하나



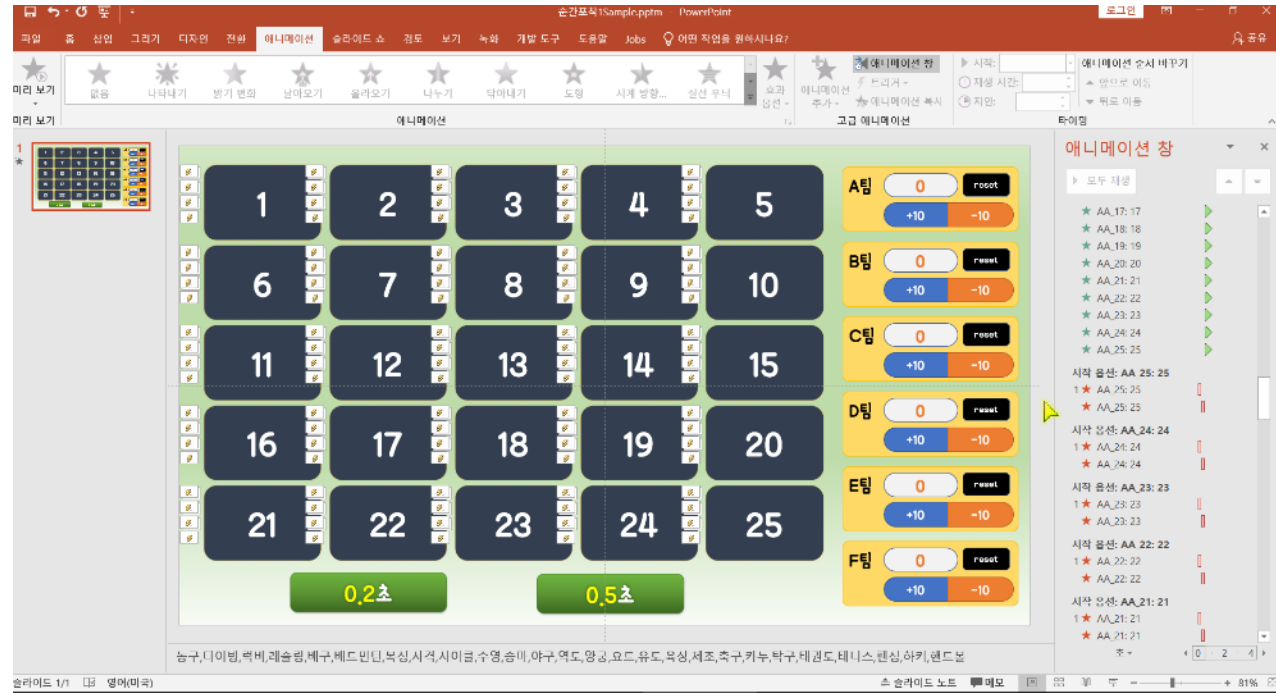
Excel VBA Macro Editor showing a script for automating data entry. The script includes comments in Korean and code for connecting to a web API, querying data, and formatting the spreadsheet.

Excel spreadsheet showing a table with columns: 날짜 (Date), 증가 (Increase), 전일비 (Previous Day's Ratio), 시가 (Opening Price), and 고가 (Closing Price). The data includes dates from 2017.02.20 to 2017.03.06 and corresponding financial values.

Bottom banner text: **Excel VBA] 매크로를 이용해 업무 자동화 하기!** | 탈잉

저작권 보호를 받는 이미지일 수 있습니다. 자세히 알아보기

방문

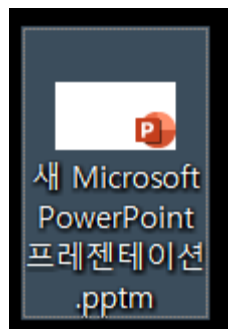


PowerPoint presentation titled '애니메이션 창' (Animation Pane). The slide displays a 5x5 grid of numbered buttons (1-25) and a control panel on the right with buttons for 'A팀', 'B팀', 'C팀', 'D팀', 'E팀', and 'F팀'. Each team button has a score display and '+10' and '-10' buttons. The bottom of the slide shows a timer with '0.2초' and '0.5초' buttons.

Animation Pane on the right lists the sequence of animation events for each element.

주로 엑셀은 사무 자동화(OA)에 많이 사용하며,
파워포인트는 게임을 만들 때 많이 사용한다.

2-5. 매크로 추출(수동)



이름

_rels
docProps
ppt

[Content_Types].xml

수정한 날짜

2021-11-04 오후 3:43

2021-11-04 오후 3:43

2021-11-04 오후 4:48

유형

파일 폴더

파일 폴더

파일 폴더

XML 문서

크기

4KB

이름

_rels
slideLayouts
slideMasters
slides
theme

presentation.xml

presProps.xml

tableStyles.xml

vbaProject.bin

viewProps.xml

수정한 날짜

2021-11-04 오후 3:43

2021-11-04 오후 3:43

2021-11-04 오후 3:43

2021-11-04 오후 3:43

2021-11-04 오후 3:43

유형

파일 폴더

파일 폴더

파일 폴더

파일 폴더

파일 폴더

XML 문서

XML 문서

XML 문서

BIN 파일

XML 문서

크기

4KB

1KB

1KB

12KB

1KB

2-5. 매크로 추출 (수동)

vbaProject.bin

```
Offset(h) 00 01 02 03 04 05 06 07
00000000 D0 CF 11 E0 A1 B1 1A E1
```

21 Results Found For D0CF11E0A1B11AE1

Extension	Signature	Description
☆ ADP	D0 CF 11 E0 A1 B1 1A E1 ASCII ••	Access project file Sizet: 8 Bytes Offset: 0 Bytes
☆ AC	D0 CF 11 E0 A1 B1 1A E1 ASCII ••	CaseWare Working Papers Sizet: 8 Bytes Offset: 0 Bytes
☆ MSI	D0 CF 11 E0 A1 B1 1A E1 ASCII ••	Microsoft Installer package Sizet: 8 Bytes Offset: 0 Bytes
☆ DB	D0 CF 11 E0 A1 B1 1A E1 ASCII ••	MSWorks database file Sizet: 8 Bytes Offset: 0 Bytes
☆ DOC	D0 CF 11 E0 A1 B1 1A E1 ASCII ••	Microsoft Office document Sizet: 8 Bytes Offset: 0 Bytes
☆ OPT	D0 CF 11 E0 A1 B1 1A E1 ASCII ••	Developer Studio File Options file Sizet: 8 Bytes Offset: 0 Bytes
☆ XLS	D0 CF 11 E0 A1 B1 1A E1 ASCII ••	Microsoft Office document Sizet: 8 Bytes Offset: 0 Bytes
☆ APR	D0 CF 11 E0 A1 B1 1A E1 ASCII ••	Lotus IBM Approach 97 file Sizet: 8 Bytes Offset: 0 Bytes
☆ MSC	D0 CF 11 E0 A1 B1 1A E1 ASCII ••	Microsoft Common Console Document Sizet: 8 Bytes Offset: 0 Bytes

이름

UserForm1
VBA
PROJECT
PROJECTwm

vbaProject - 복사본 > VBA

이름




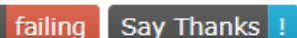
_VBA_PROJECT
dir
Module1
UserForm1

Module1

```
Offset(h) 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F Decoded text
00000580 00 6F 00 FF FF 20 00 00 00 FF FF FF FF B8 00 00 .o.ÿÿ ...ÿÿÿÿ...
00000590 00 FF FF FF FF 00 00 01 9E B0 00 41 74 74 72 69 .ÿÿÿÿ...ž°.Attri
000005A0 62 75 74 00 65 20 56 42 5F 4E 61 6D 00 65 20 3D but.e VB_Nam.e =
000005B0 20 22 4D 6F 64 00 75 6C 65 31 22 0D 0A 53 00 75 "Mod.ule1"..S.u
000005C0 62 20 61 75 74 6F 5F 00 6F 70 65 6E 28 29 0D 0A b auto_.open()..
000005D0 00 4D 73 67 42 6F 78 20 22 00 B8 C5 C5 A9 B7 CE .MsgBox "...AA@·f
000005E0 20 BD 08 C7 C7 E0 00 4C 45 6E 64 20 09 00 54 0D ¼.ÇÇà.LEnd ..T.
000005F0 0A 08 62 63 6C 6F 73 0A 65 01 64 20 00 00 41 70 ..bclos.e.d ..Ap
00000600 70 6C 00 69 63 61 74 69 6F 6E 2E 00 43 6F 6D 6D plication..Comm
00000610 61 6E 64 42 00 61 72 73 28 22 74 6F 6F 00 6C 73 andB.ars("too.ls
00000620 22 29 2E 52 65 73 10 65 74 0D 0A 0D 48 73 64 66 ") .Res.et...Hsdf
00000630 21 09 75 74 65 73 74 09 6E !.utest.n
```

2-6. 매크로 추출 (OLEVBA 사용)

python-oletools

`oletools` is a package of python tools to analyze Microsoft (File Binary Format or Compound Document File Format), su messages, mainly for malware analysis, forensics and debug <http://www.decalage.info/python/oletools> for more info.

Quick links: [Home page](#) - [Download/Install](#) - [Documentatic](#)
[Author](#) - [Repository](#) - [Updates on Twitter](#) [Cheatsheet](#)

Note: python-oletools is not related to OLETools published

Supported formats

- Word 97-2003 (.doc, .dot), Word 2007+ (.docm, .dotm)
- Excel 97-2003 (.xls), Excel 2007+ (.xlsm, .xlsb)
- PowerPoint 97-2003 (.ppt), PowerPoint 2007+ (.pptm, .ppsm)
- Word/PowerPoint 2007+ XML (aka Flat OPC)
- Word 2003 XML (.xml)
- Word/Excel Single File Web Page / MHTML (.mht)
- Publisher (.pub)
- SYLK/SLK files (.slk)
- Text file containing VBA or VBScript source code
- Password-protected Zip archive containing any of the above

2-6. 매크로 추출 (OLEVBA 사용)

Windows

```
pip install -U https://github.com/decalage2/oletools/archive/master.zip
```

```
>olevba 1234.pptm
```

```
-----  
VBA MACRO Module1  
in file: ppt/vbaProject.bin - OLE stream: 'Module1'  
-----  
Sub auto_open()  
MsgBox " , ÅÅ@ · ↑ ½ÇÇà"  
End Sub  
  
Sub auto_close()  
Application.CommandBars("tools").Reset  
End Sub  
  
Sub asdf()  
MsgBox "test"  
End Sub  
-----
```

2-7. 진단 원리

```
Private Sub fngjksnhokdnfd(newValue As Integer)
Dim wsh As Object
Dim regKey As String
Set wsh = CreateObject("WScript.shell")
regKey = "HKEY_CURRENT_USER\Software\Microsoft\Office\" & Application.Version &
"\Word\Security\AccessVBOM"
wsh.RegWrite regKey, newValue, "REG_DWORD"
End Sub
```

VBA 문법은 VB6의 문법과 동일하며,
VB6 특성상, 호출되는 함수명이 반드시 노출되는 점이 발생

즉, 코드 내부 VB 내장 함수와 Win32API 함수가
반드시 드러나는 구조이므로 이를 통해 진단

2-8. 진단방안

```
def MalFunc(Funcname):  
    # 1. 파일 접근 관련 (R/W)  
    if Funcname.find("For Output As")!=-1: return "파일 접근 관련 (Open문 Output)"  
    if Funcname.find("For Input As")!=-1: return "파일 접근 관련 (Open문 Input)"  
    if Funcname.find("For Append As")!=-1: return "파일 접근 관련 (Open문 Appen)"  
    if Funcname.find("URLDownloadToFile")!=-1: return "파일 접근 관련 (URLDownloadToFile API)"  
    if Funcname.find("WriteFile")!=-1: return "파일 접근 관련 (WriteFile API)"  
    if Funcname.find("ReadFile")!=-1: return "파일 접근 관련 (ReadFile API)"  
    if Funcname.find("FindFirstFile")!=-1: return "파일 검색 관련 (FindFirstFile API)"  
    if Funcname.find("FileCopy(")!=-1: return "파일 복사 관련 (FileCopy 함수)"  
    if Funcname.find("FileCopy ")!=-1: return "파일 복사 관련 (FileCopy 함수)"  
    if Funcname.find("CopyFile")!=-1: return "파일 복사 관련 (CopyFile API)"  
    if Funcname.find("Kill ")!=-1: return "파일 삭제 관련 (Kill 함수)"  
    if Funcname.find("Kill(")!=-1: return "파일 삭제 관련 (Kill 함수)"  
    if Funcname.find("DeleteFile")!=-1: return "파일 삭제 관련 (DeleteFile API)"  
    if Funcname.find("CreateFile")!=-1: return "파일 삭제 관련 (CreateFile API)"
```

2-8. 진단방안

2. 파일 실행 관련

```
if Funcname.find("ShellExecute")!=-1: return "파일 실행 관련 (ShellExecut API)"
```

```
if Funcname.find("Shell ")!=-1: return "파일 실행 관련 (Shell 함수)"
```

```
if Funcname.find("Shell(")!=-1: return "파일 실행 관련 (Shell 함수)"
```

3. 오브젝트 관련

```
if Funcname.find("GetObject")!=-1: return "오브젝트 관련 (GetObject API)"
```

```
if Funcname.find("CreateObject")!=-1: return "오브젝트 관련 (CreateObject API)"
```

```
if Funcname.find("As Object")!=-1: return "오브젝트 변수"
```

4. 메모리 관련

```
if Funcname.find("ReadProcessMemory")!=-1: return "메모리 접근 관련 (ReadProcessMemory API)"
```

```
if Funcname.find("WriteProcessMemory")!=-1: return "메모리 접근 관련 (WriteProcessMemory API)"
```

5. 프로세스 관련

```
if Funcname.find("CreateProcess")!=-1: return "프로세스 관련 (CreateProcess API)"
```

```
if Funcname.find("OpenProcess")!=-1: return "프로세스 관련 (OpenProcess API)"
```

```
if Funcname.find("WinExec")!=-1: return "프로세스 관련 (WinExec API)"
```

```
if Funcname.find("SendMessage")!=-1: return "프로세스 관련 (SendMessage API)"
```

```
if Funcname.find("PostMessage")!=-1: return "프로세스 관련 (PostMessage API)"
```

2-8. 진단방안

6. 레지스트리 관련

```
if Funcname.find("RegSetValue")!=-1: return "레지스트리 관련 (RegSetValue API)"
if Funcname.find("RegEnumValue")!=-1: return "레지스트리 관련 (RegEnumValue API)"
if Funcname.find("RegDeleteKey")!=-1: return "레지스트리 관련 (RegDeleteKey API)"
if Funcname.find("RegCreateKey")!=-1: return "레지스트리 관련 (RegCreateKey API)"
if Funcname.find("OpenProcessToken")!=-1: return "레지스트리 관련 (OpenProcessToken API)"
```

7. 마우스, 키보드 제어 관련

```
if Funcname.find("mouse_event")!=-1: return "마우스 제어 관련 (mouse_event API)"
if Funcname.find("keybd_event")!=-1: return "키보드 제어 관련 (keybd_event API)"
if Funcname.find("SetCursorPos")!=-1: return "마우스 제어 관련 (SetCursorPos API)"
```

8. 스레드 관련

```
if Funcname.find("CreateThread")!=-1: return "스레드 관련 (CreateThread API)"
if Funcname.find("ResumeThread")!=-1: return "스레드 관련 (ResumeThread API)"
if Funcname.find("CreateRemoteThread")!=-1: return "스레드 관련 (CreateRemoteThread API)"
```

9. 정상 파일

```
return 1
```

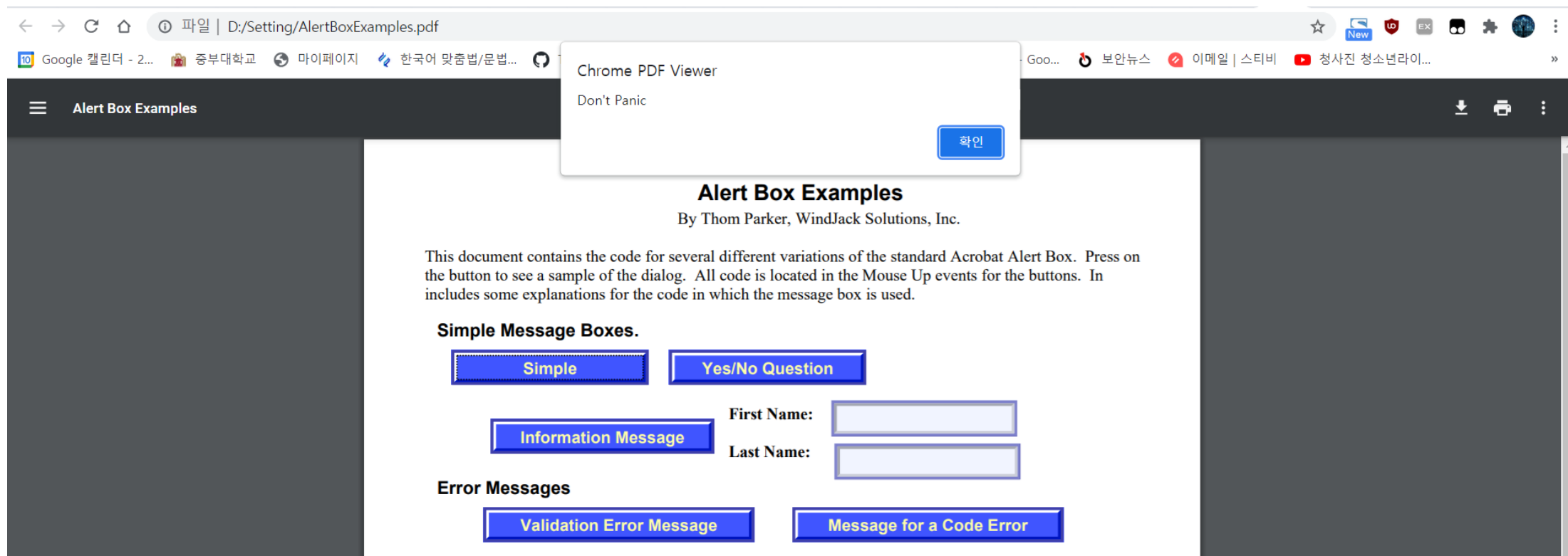

2-8. 진단방안

```
if os.path.isfile(path1) :
    with open(path1, "rt", encoding='UTF8') as fFILE :
        Cou = 0
        while True :
            line = fFILE.readline()
            if not line :
                if (MalCou == 1):
                    print("※ 악성 매크로 파일")
                else:
                    print("※ 정상 매크로 파일")
                break
            Cou = Cou + 1
            if MalFunc(line) != 1 :
                print("악성 명령 감지 [line :", Cou, ", ", MalFunc(line), "]")
                MalCou = 1
```

3. PDF 악성 문서 진단

3-1. PDF 악성 문서

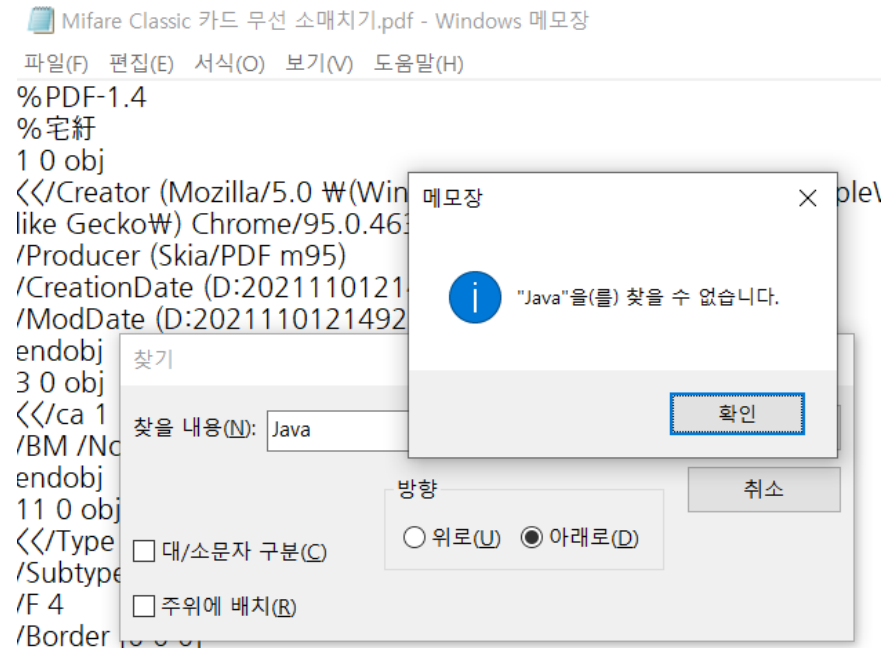
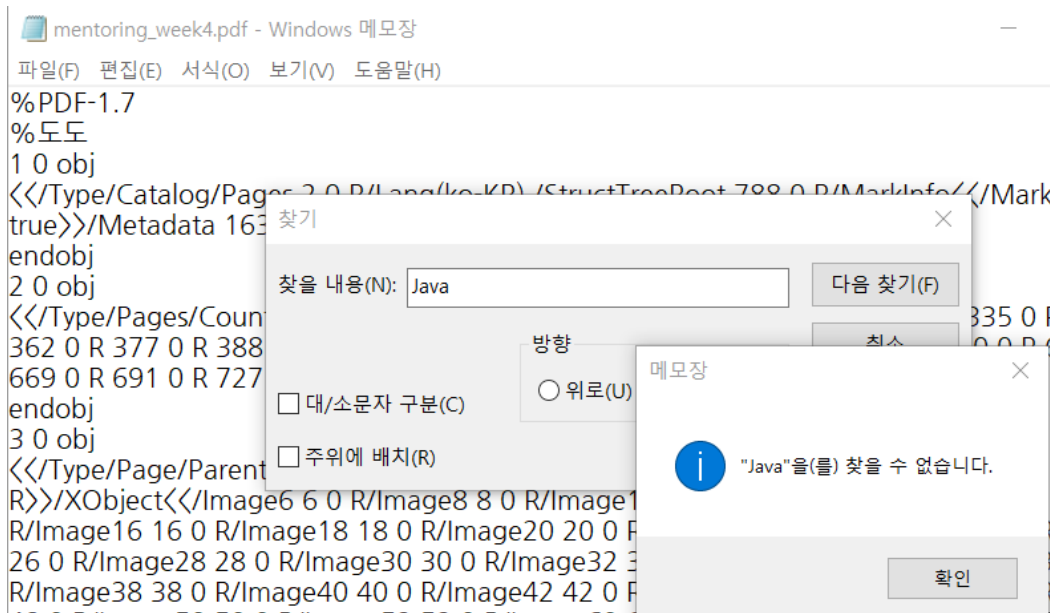
PDF는 JavaScript를 지원하고, 이를 통해 악성 행위가 가능함.



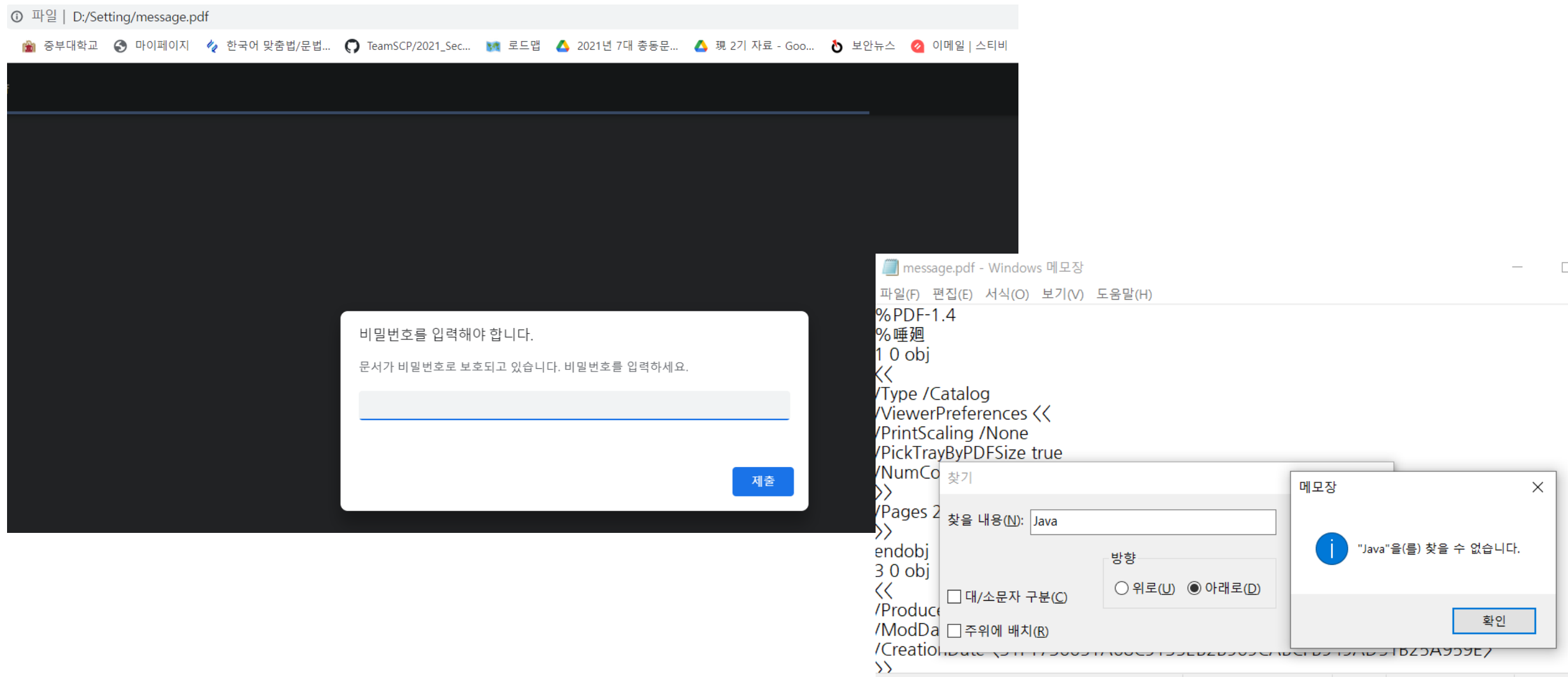
3-2. 진단 원리

JS는 난독화가 용이해서 조금 더 근본적으로 접근해서 진단.


과연 일반적인 PDF는 JS를 사용할까?



3-2. 진단 원리



3-3. 진단방안

 AlertBoxExamples.pdf - Windows 메모장

파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)

82 0 obj

<</Subtype/TrueType/FontDescriptor 104 0 R/LastChar 32/Widths
[278]/BaseFont/ArialMT/FirstChar 32/Encoding/WinAnsiEncoding/Type/Font>>

endobj

83 0 obj

<</S/JavaScript/JS(app.alertW("Don't Panic"~);WrWn)>>

endobj

84 0 obj

<</S/JavaScript/JS 85 0 R>>

JS를 사용하는 PDF 문서의 특징

3-3. 진단방안

1. 파일을 한줄 한줄 EOF까지 읽음.

1-1. 각각의 라인에 << 와 >> 문자가 포함되어 있는지 체크

1-1-1. 포함되어 있으면 문자를 전부 소문자로 치환

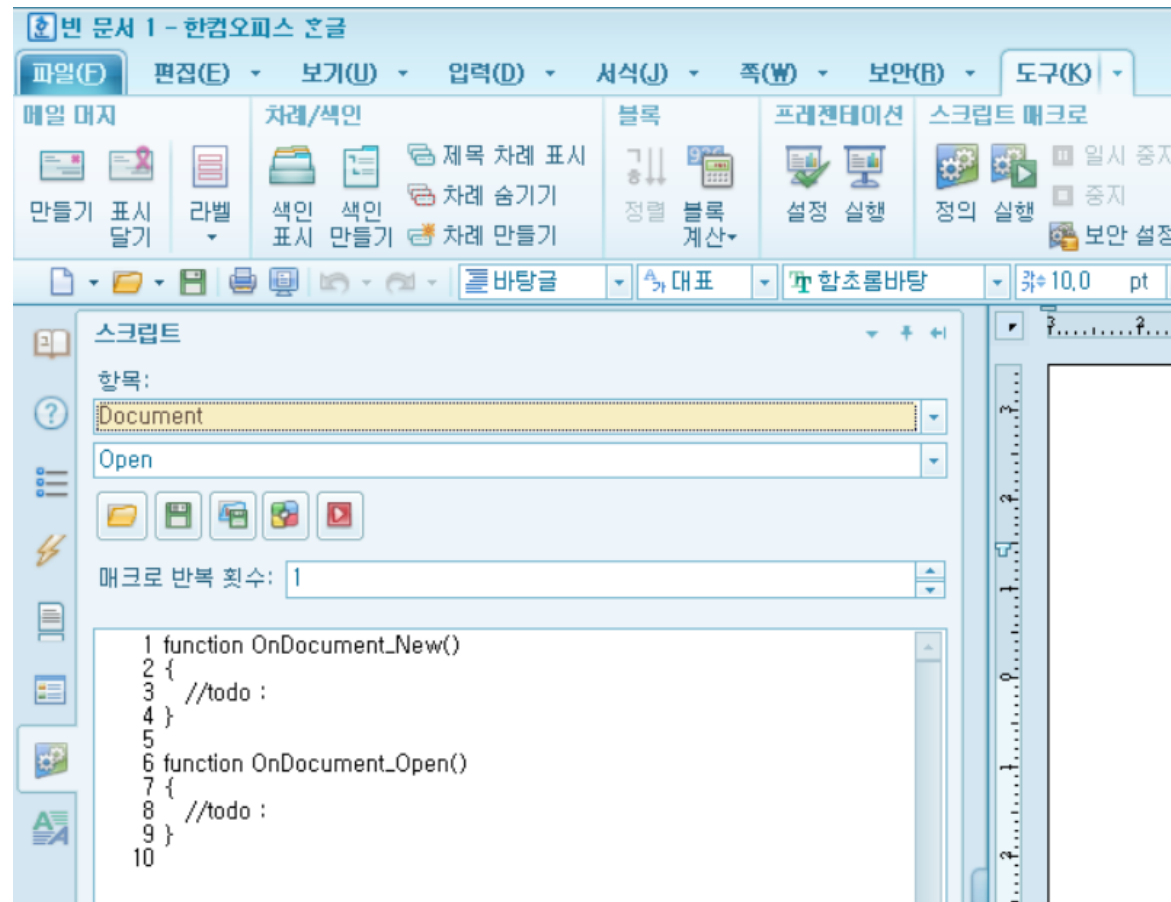
1-1-2. 라인에 /javascript 또는 /js 문구가 들어가있는지 체크

1-1-2-1. 들어가 있으면 악성 문서로 진단

2. EOF까지 읽었을 때 자바스크립트 문구가 없으면 정상 문서로 진단

4. 한글(HWP) 악성 문서 진단

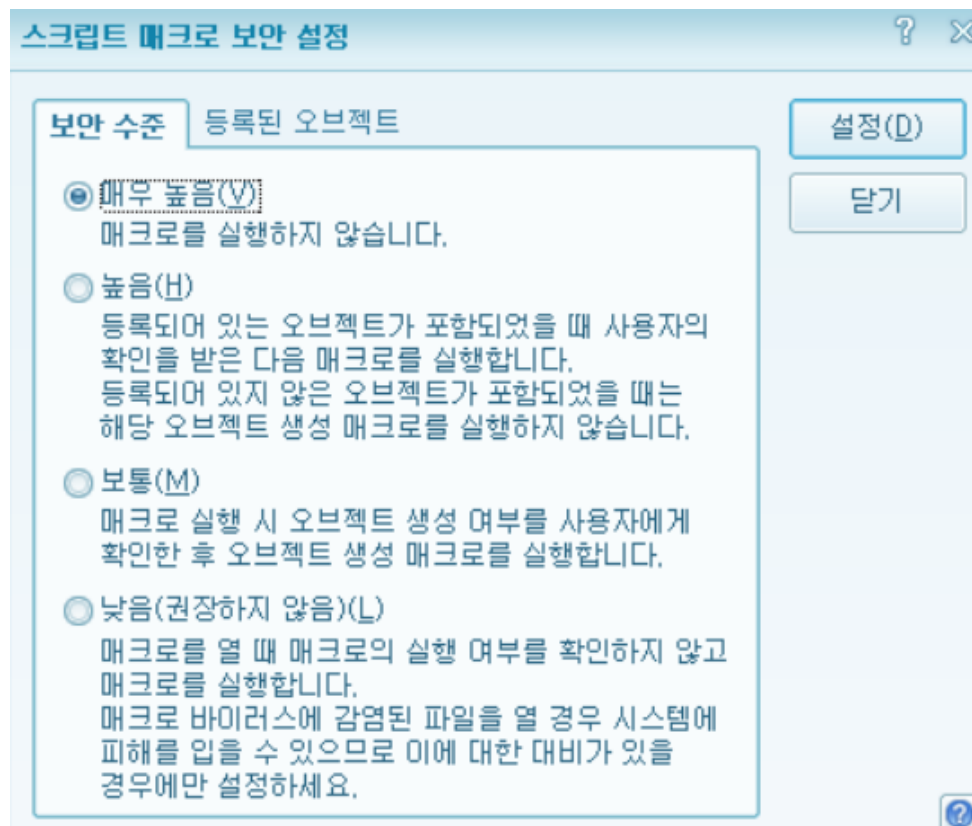
4-1. 아래 한글의 매크로



아래 한글의 매크로는 PostScript, GhostScript 라는 스크립트를 사용함.

3-2. 진단 원리

과연 일반적인 한글 파일은 JS를 사용할까?



3-3. 진단 방안 : hwp 5.x

설명	구별 이름	길이(바이트)	레코드 구조	압축/암호화
파일 인식 정보	FileHeader	고정		
문서 정보	DocInfo	고정	✓	✓
본문	BodyText Section0 Section1 ...	가변	✓	✓
문서 요약	\005HwpSummaryInformation	고정		
바이너리 데이터	BinData BinaryData0 BinaryData1 ...	가변		✓
미리보기 텍스트	PrvText	고정		
미리보기 이미지	PrvImage	가변		
문서 옵션	DocOptions _LinkDoc DrmLicense ...	가변		
스크립트	Scripts DefaultJScript JScriptVersion ...	가변		
XML 템플릿	XMLTemplate Schema Instance ...	가변		
문서 이력 관리	DocHistory VersionLog0 VersionLog1 ...	가변	✓	✓

hwp 전체 구조

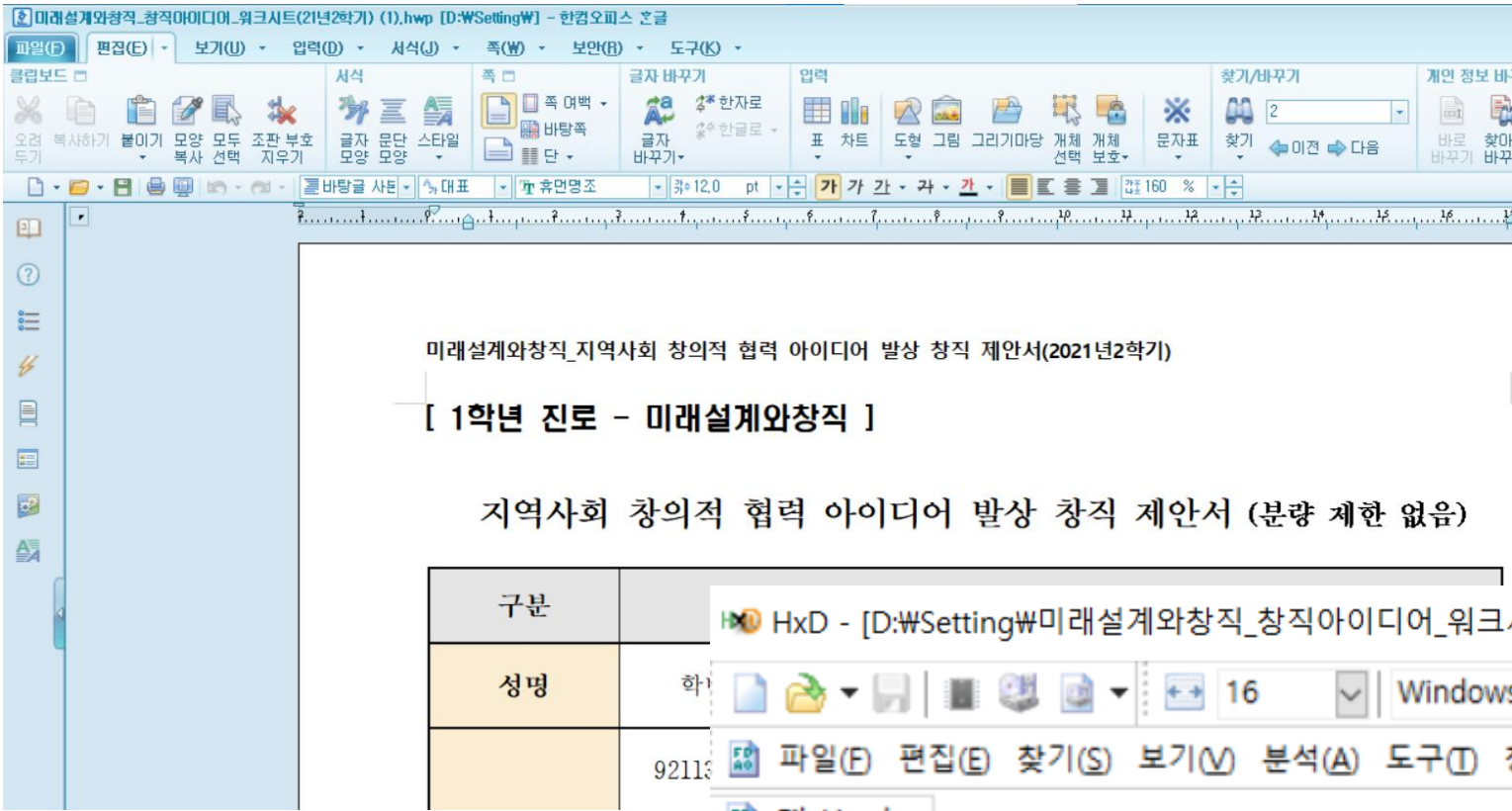
HWP 5.x 버전은
OLE 구조를 띄고 있음.

3-3. 진단 방안 : hwp 5.x

자료형	길이(바이트)	설명
BYTE array[32]	32	signature. 문서 파일은 "HWP Document File"
DWORD	4	<p>파일 버전. 0xMMnnPPrr의 형태(예 5.0.3.0)</p> <ul style="list-style-type: none"> MM: 문서 형식의 구조가 완전히 바뀌는 것을 나타냄. 숫자가 다르면 구 버전과 호환 불가능. nn: 큰 구조는 동일하나, 큰 변화가 있는 것을 나타냄. 숫자가 다르면 구 버전과 호환 불가능. PP: 구조는 동일, Record가 추가되었거나, 하위 버전에서 호환되지 않는 정보가 추가된 것을 나타냄. 숫자가 달라도 구 버전과 호환 가능. rr: Record에 정보들이 추가된 것을 나타냄. 숫자가 달라도 구 버전과 호환 가능.
DWORD	4	속성
		범위
		설명
		bit 0 압축 여부
		bit 1 암호 설정 여부
		bit 2 배포용 문서 여부
		bit 3 스크립트 저장 여부
		bit 4 DRM 보안 문서 여부
		bit 5 XMLTemplate 스토리지 존재 여부
		bit 6 문서 이력 관리 존재 여부
		bit 7 전자 서명 정보 존재 여부
		bit 8 공인 인증서 암호화 여부
		bit 9 전자 서명 예비 저장 여부
		bit 10 공인 인증서 DRM 보안 문서 여부
		bit 11 CCL 문서 여부
		bit 12 모바일 최적화 여부
		bit 13 개인 정보 보안 문서 여부
		bit 14 변경 추적 문서 여부
		bit 15 공공누리(KOGL) 저작권 문서
		bit 16 비디오 컨트롤 포함 여부
		bit 17 차례 필드 컨트롤 포함 여부
		bit 18~31 예약

파일 헤더 정보의 속성 태그에
스크립트 저장 여부를
비트로 저장하고 있음.

3-3. 진단 방안 : hwp 5.x



일반 파일의 경우

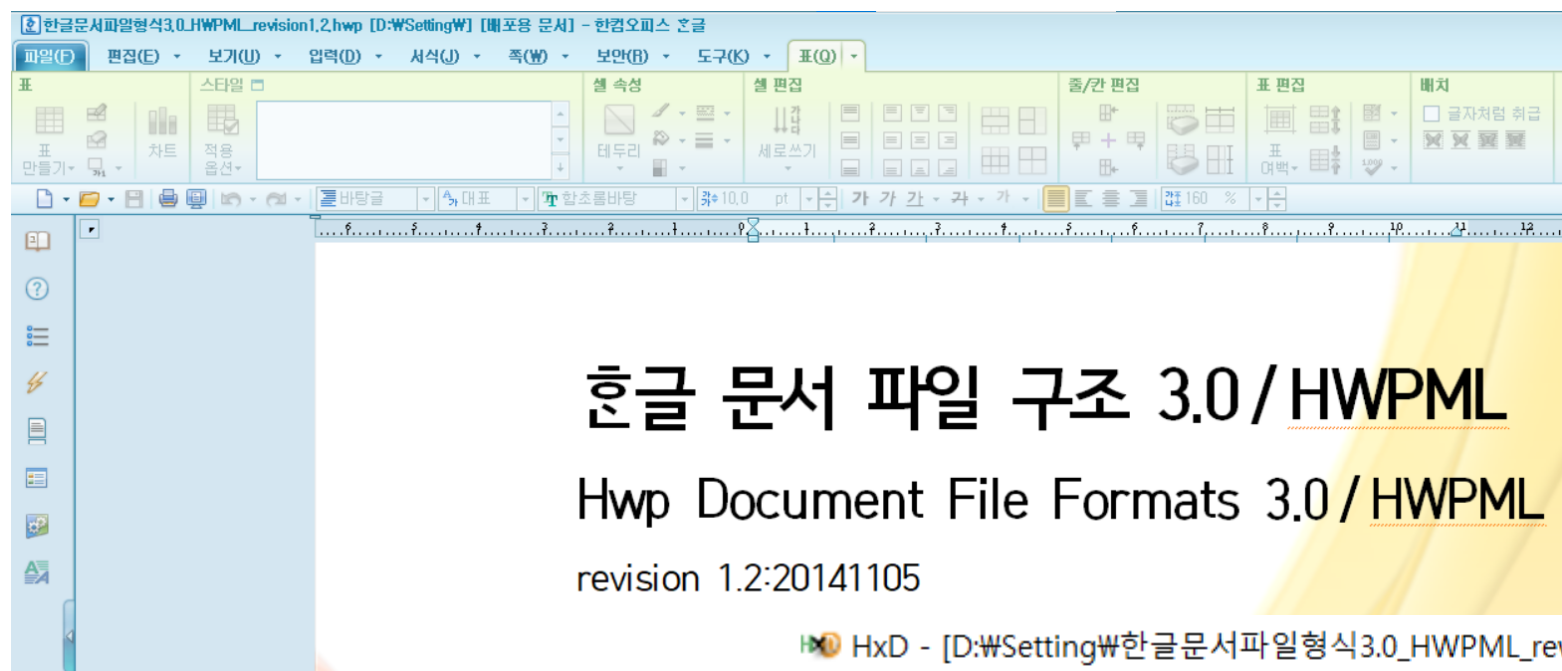
Offset (h)	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	Decoded text
00000000	48	57	50	20	44	6F	63	75	6D	65	6E	74	20	46	69	6C	HWP Document Fil
00000010	65	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	e.....
00000020	00	03	00	05	01	00	00	00	00	00	00	00	04	00	00	00
00000030	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00

3-3. 진단 방안 : hwp 5.x

HEX	1
DEC	1
OCT	1
BIN	0001

속성	
범위	설명
bit 0	압축 여부
bit 1	암호 설정 여부
bit 2	배포용 문서 여부
bit 3	스크립트 저장 여부
bit 4	DRM 보안 문서 여부
bit 5	XMLTemplate 스토리지 존재 여부
bit 6	문서 이력 관리 존재 여부
bit 7	전자 서명 정보 존재 여부
bit 8	공인 인증서 암호화 여부
bit 9	전자 서명 예비 저장 여부
bit 10	공인 인증서 DRM 보안 문서 여부
bit 11	CCL 문서 여부
bit 12	모바일 최적화 여부
bit 13	개인 정보 보안 문서 여부
bit 14	변경 추적 문서 여부
bit 15	공공누리(KOGL) 저작권 문서
bit 16	비디오 컨트롤 포함 여부
bit 17	차례 필드 컨트롤 포함 여부
bit 18 ~ 31	예약

3-3. 진단 방안 : hwp 5.x



배포용 문서의 경우

HxD - [D:\SettingW\한글문서파일형식3.0_HWPML_revision1.2 - 복사본\FileHeader]

FileHeader

Offset(h)	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	Decoded text
00000000	48	57	50	20	44	6F	63	75	6D	65	6E	74	20	46	69	6C	HWP Document Fil
00000010	65	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	e.....
00000020	00	05	00	05	05	00	00	00	00	00	00	00	04	00	00	00

3-3. 진단 방안 : hwp 5.x

HEX 5
 DEC 5
 OCT 5
 BIN 0101

속성	
범위	설명
bit 0	압축 여부
bit 1	암호 설정 여부
bit 2	배포용 문서 여부
bit 3	스크립트 저장 여부
bit 4	DRM 보안 문서 여부
bit 5	XMLTemplate 스토리지 존재 여부
bit 6	문서 이력 관리 존재 여부
bit 7	전자 서명 정보 존재 여부
bit 8	공인 인증서 암호화 여부
bit 9	전자 서명 예비 저장 여부
bit 10	공인 인증서 DRM 보안 문서 여부
bit 11	CCL 문서 여부
bit 12	모바일 최적화 여부
bit 13	개인 정보 보안 문서 여부
bit 14	변경 추적 문서 여부
bit 15	공공누리(KOGL) 저작권 문서
bit 16	비디오 컨트롤 포함 여부
bit 17	차례 필드 컨트롤 포함 여부
bit 18 ~ 31	예약

3-3. 진단 방안: hwp 5.x

스크립트

항목: Document

Open

매크로 반복 횟수: 1

```

1 function OnDocument_New()
2 {
3     //todo :
4 }
5
6 function OnDocument_Open()
7 {
8     //todo :
9
10    HParameterSet, HInsertText, Text = "Hello World!";
11    HAction, Execute("InsertText", HParameterset, HInsertText, Hset);
12 }
13

```

Hello World!!

스크립트가 저장된 경우

HxD - [C:\Users\W2NW\Desktop\스크립트 - 복사본\FileHeader]

16 Windows (ANSI) 16진수

파일(F) 편집(E) 찾기(S) 보기(V) 분석(A) 도구(T) 창 설정(W) 도움말(H)

FileHeader

Offset (h)	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	Decoded text
00000000	48	57	50	20	44	6F	63	75	6D	65	6E	74	20	46	69	6C	HWP Document Fil
00000010	65	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	e.....
00000020	00	03	00	05	09	00	00	00	00	00	00	00	04	00	00	00

3-3. 진단 방안 : hwp 5.x

HEX	9
DEC	9
OCT	11
BIN	1001

속성	
범위	설명
bit 0	압축 여부
bit 1	암호 설정 여부
bit 2	배포용 문서 여부
bit 3	스크립트 저장 여부
bit 4	DRM 보안 문서 여부
bit 5	XMLTemplate 스토리지 존재 여부
bit 6	문서 이력 관리 존재 여부
bit 7	전자 서명 정보 존재 여부
bit 8	공인 인증서 암호화 여부
bit 9	전자 서명 예비 저장 여부
bit 10	공인 인증서 DRM 보안 문서 여부
bit 11	CCL 문서 여부
bit 12	모바일 최적화 여부
bit 13	개인 정보 보안 문서 여부
bit 14	변경 추적 문서 여부
bit 15	공공누리(KOGL) 저작권 문서
bit 16	비디오 컨트롤 포함 여부
bit 17	차례 필드 컨트롤 포함 여부
bit 18 ~ 31	예약

3-3. 진단방안: hwp 5.x

1. HWP 5.0 기반 한글 문서의 경우

1-1. 파일 헤더 정보의 속성 태그 비트를 검증함

1-2. 세 번째 비트 활성화시, 악성 문서로 진단.

비활성화시, 정상 문서로 진단.

3-4. 진단방안: HWPML

HWPML 파일은 HWP의 개방형 문서 형식으로,
마크업 언어를 기반으로 하고 있는 문서이다.

HxD - [C:\Users\W2NW\Desktop\스크립트.html]

Windows (ANSI) 16진수

파일(F) 편집(E) 찾기(S) 보기(V) 분석(A) 도구(T) 창 설정(W) 도움말(H)

스크립트.html

Offset (h)	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	Decoded text
00000000	EF	BB	BF	3C	3F	78	6D	6C	20	76	65	72	73	69	6F	6E	i»¿<?xml version
00000010	3D	22	31	2E	30	22	20	65	6E	63	6F	64	69	6E	67	3D	= "1.0" encoding=
00000020	22	55	54	46	2D	38	22	20	73	74	61	6E	64	61	6C	6F	"UTF-8" standalo
00000030	6E	65	3D	22	6E	6F	22	20	3F	3E	3C	48	57	50	4D	4C	ne="no" ?><HWPML
00000040	20	53	74	79	6C	65	3D	22	65	6D	62	65	64	22	20	53	Style="embed" S
00000050	75	62	56	65	72	73	69	6F	6E	3D	22	38	2E	30	2E	30	ubVersion="8.0.0
00000060	2E	30	22	20	56	65	72	73	69	6F	6E	3D	22	32	2E	38	.0" Version="2.8
00000070	22	3E	3C	48	45	41	44	20	53	65	63	43	6E	74	3D	22	"><HEAD SecCnt="
00000080	31	22	3E	3C	44	4F	43	53	55	4D	4D	41	52	59	3E	3C	1"><DOCSUMMARY><
00000090	41	55	54	48	4F	52	3E	32	4E	3C	2F	41	55	54	48	4F	AUTHOR>2N</AUTHO
000000A0	52	3E	3C	44	41	54	45	3E	32	30	32	31	EB	85	84	20	R><DATE>2021ë....,
000000B0	31	30	EC	9B	94	20	32	38	EC	9D	BC	20	EB	AA	A9	EC	10ì>" 28ì.¼ ë*@ì
000000C0	9A	94	EC	9D	BC	20	EC	98	A4	ED	9B	84	20	32	3A	33	š"ì.¼ ì"¼ì>,, 2:3
000000D0	36	3A	32	34	3C	2F	44	41	54	45	3E	3C	2F	44	4F	43	6:24</DATE></DOC
000000E0	53	55	4D	4D	41	52	59	3E	3C	4F	43	53	45	54	54		SUMMARY><DOCSETT
000000F0	49	4E	47	3E	3C	42	45	47	49	4E	4E	55	4D	42	45	52	ING><BEGINNUMBER

3-4. 진단방안: HWPML

[illegible]

<SCRIPTCODE>

<SCRIPTHEADER>

<SCRIPTSOURCE>

헤더가 기본적으로 붙는걸 알 수 있다.

3-4. 진단방안: HWPML

2. HWPML 기반 한글 문서의 경우

2-1. 스크립트 태그를 사용한다면

2-1-1. 디폴트 스크립트면 정상 문서로
아니면 악성 문서로 간주

2-2. 스크립트 태그를 사용하지 않으면 정상 문서로 간주

Q & A

「
감사합니다
」