

악성문서 분석 요점 정리

일반적인 접근 방법론

1. 셸코드, VBA 매크로, 자바스크립트 등 문서에 포함된 모든 코드의 위치 파악
2. 의심스러운 코드 영역을 추출
3. 필요하다면, 셸코드를 디스어셈블하거나 디버깅
4. 필요하다면 난독화된 자바스크립트, 액션스크립트, VB 매크로 코드를 분석
5. 감염 경로의 다음 단계를 이해

MS 오피스 바이너리 파일 포맷

구조화된 저장소 (OLE SS)는 바이너리 MS 오피스 파일 내부의 파일 시스템을 정의

데이터는 "스토리지"(폴더)이거나 "스트림" (파일)

엑셀은 "workbook" 스트림 내부에 데이터 저장

파워포인트는 "PowerPoint Document" 스트림 내부에 데이터 저장

워드는 다양한 스트림 내부에 데이터 저장

MS 오피스 파일 분석 도구 목록

OfficeMalScanner 는 MS 오피스 파일 (DOC, XLS, PPT) 파일의 셸코드와 VBA 매크로 위치를 검색

DisView 는 MS 오피스 파일의 오프셋을 입력으로 받아 디스어셈블을 수행 (OfficeMalScanner 의 일부)

MalHost-Setup 은 MS 오피스 파일의 특정 오프셋에서 셸코드를 추출하여 EXE 생성 (OfficeMalScanner 의 일부)

Offvis 는 MS 오피스 파일의 내용과 구조를 있는 그대로 보여주고 몇몇 익스플로잇을 찾아내는 기능을 가지고 있음

BIFF-Workbench 는 XLS 파일의 원본 내용과 구조를 보여주고 편집과 검색 기능을 제공

Office Binary Translator 는 DOC, PPT, XLS 파일을 Open XML 파일로 변환 (BiffView 도구 포함)

OfficeCat 는 MS 오피스 파일에서 이미 알려진 취약점을 공격하는 익스플로잇을 검사

유용한 MS 오피스분석 명령 목록

OfficeMalScanner *file.doc* 내부의 셸코드, OLE *file.doc scan brute* 데이터, PE 파일 검색

OfficeMalScanner *file.doc* 의 VB 매크로 탐색 *file.doc info*

OfficeMalScanner *file.docx* 압축을 해제한 후 VB 매크로 코드 탐색 *file.docx inflate*

DisView *file.doc* *file.doc* 의 0x4500 위치의 셸코드 디스어셈블 *0x4500*

MalHost-Setup *file.doc out.exe* *file.doc* 0x4500 오프셋의 셸코드를 *out.exe* 파일로 생성 *0x4500*

어도비 PDF 파일 포맷

PDF 파일은 헤더, 개체, 크로스-레퍼런스 테이블 (개체 위치 탐색용), 트레일러로 구성

"/OpenAction"과 "/AA" (Additional Action)은 자동으로 실행될 스크립트나 액션을 지정

"/Names", "/AcroForm", "/Action"도 스크립트나 액션 실행

"/JavaScript"은 실행될 자바스크립트를 지정

"/GoTo*"는 같은 파일이나 다른 파일의 특정 위치로 뷰 전환

"/Launch"는 프로그램을 실행하거나 문서 열기 수행

"/URI"는 해당 URL 이 지시하는 리소스에 액세스

"/SubmitForm"과 "/GoToR"은 URL 로 데이터 전송 가능

"/RichMedia"은 PDF 에 플래시를 포함할 때 사용될 수 있음

"/ObjStm"은 개체 스트림 내부에 개체를 숨길 수 있음

"/JavaScript" 를 "/J#61vaScript"처럼 16 진수로 난독화 하는 경우를 주의하여 분석할 것 (예제 참조)

어도비 PDF 파일 분석 도구 목록

PDF StructAzer 는 PDF 파일의 구조와 원본 내용을 보거나 편집할 수 있음 (사용자 매뉴얼 참조)

PDFid 는 스크립트나 액션과 관련된 문자열을 담은 PDF 들을 식별 (파이썬 PDF Tools 의 일부)

PDF-parser 는 PDF 를 렌더링 하지 않고도 주요 요소들을 식별할 수 있음 (파이썬 PDF Tools 의 일부)

Origami 는 PDF 파일을 파싱, 분석, 편집, 생성하는 루비 프레임워크

Sumatra PDF 와 MuPDF 는 어도비 아크로벳 대응으로 쓸 수 있는 무료 PDF 뷰어

Pdftk 는 PDF 를 조작하거나 페이지 스트림 압축 해제 가능

Malzilla 는 PDF 내의 zlib 스트림을 압축 해제할 수 있고 난독화된 자바스크립트 해석에 도움이 될 수 있음

Jsunpack-n 은 PCAP 파일에서 자바스크립트를 추출 및 디코드할 수 있고 PDF 파일을 디코드 할 수 있음

CWSandbox 와 Wepawet 는 악성 PDF 파일의 일부 특성들을 자동으로 분석하는데 사용할 수 있음

유용한 PDF 분석 명령 목록

pdfid.py *file.pdf* *file.pdf* 내의 스크립트나 액션 관련 문자열 탐색

pdf-parser.py *file.pdf* *file.pdf* 의 구조를 보여주므로 의심스러운 요소 파악 가능

pdfscan.rb *file.pdf* *file.pdf* 의 구조 표시 (사용법)

pdftk *file.pdf* *file.pdf* 의 페이지 스트림 압축을 해제하여 *out.pdf* 파일로 저장 *output out.pdf* *uncompress*

악성 파일 분석 도구 목록

McAfee FileInsight 는 16 진수 편집기, 계산기, 디스어셈블러, 디코더, 스크립팅 지원 기능을 통합 제공

ExeFilter 는 오피스와 PDF 파일의 스크립트를 필터링 가능

VirusTotal 은 다양한 AV 로 파일을 스캔하여 일부 악성 문서를 인식할 수 있음

레퍼런스

어도비 포터블 도큐먼트 포맷 (PDF) 레퍼런스

PDF 파일의 물리적 구조와 논리적 구조

오피스 파일을 이용한 표적 공격 분석 (비디오)

OfficeMalScanner 이용한 MS 오피스 맬웨어 분석 (추가 자료)

PDF 보안 분석과 맬웨어의 위험

PDF 내의 악성 오리가미 (추가 프레젠테이션 자료)

OffVis 1.0 Beta: 오피스 문서 시각화 도구

리버스 엔지니어링 요점 정리