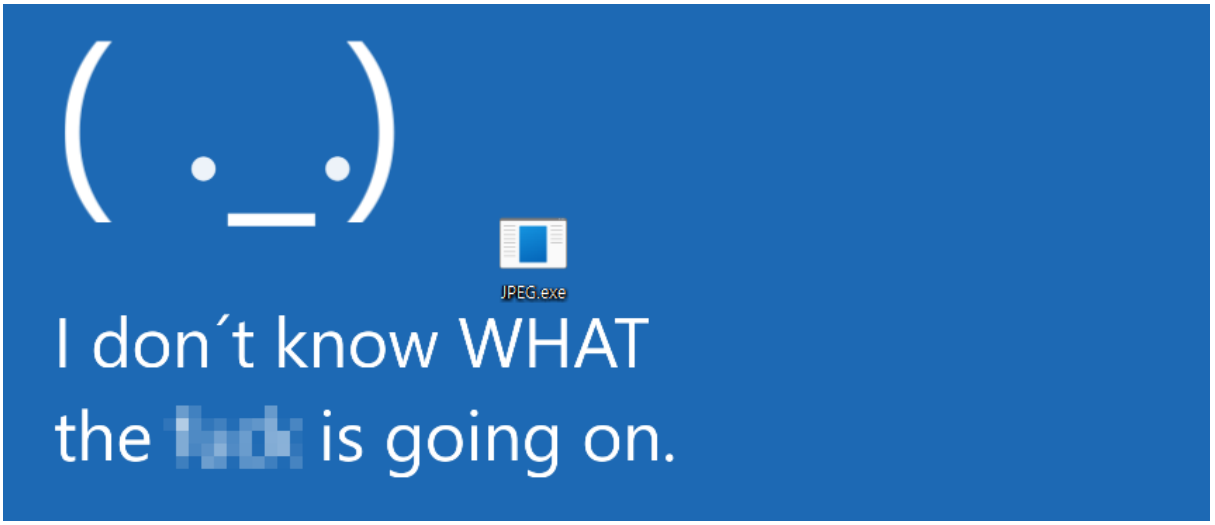
Do_you_know_jpeg?

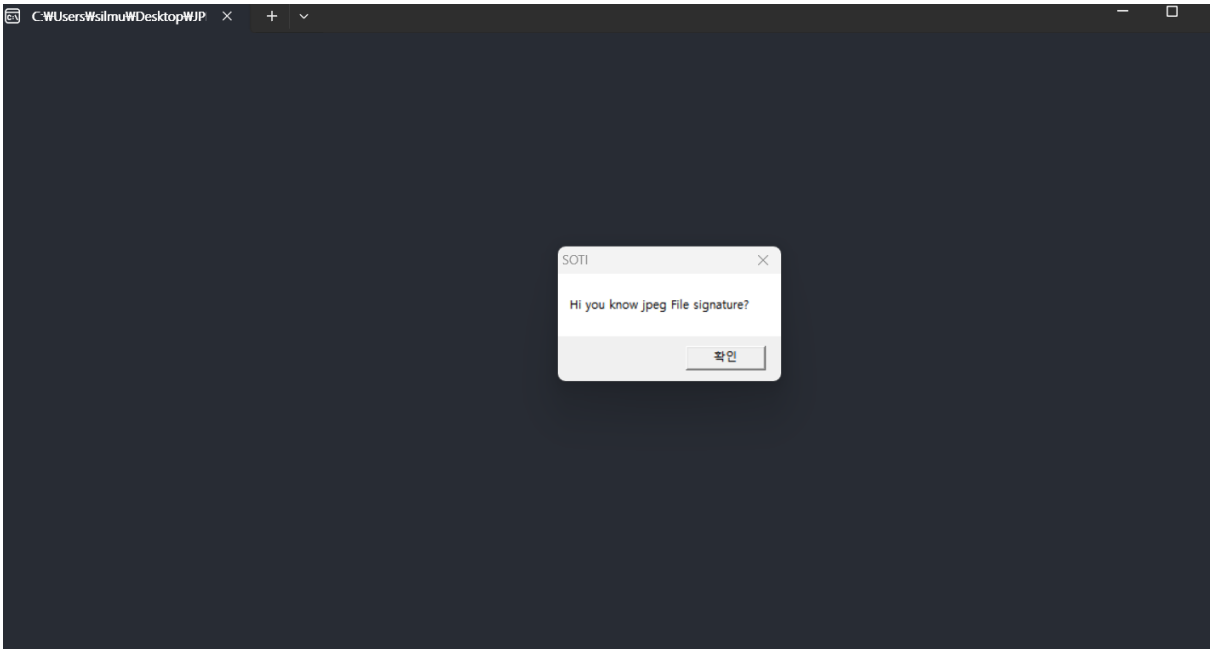
문제 분류	forensic
poc 작성자	김재환
문제 개발자	최보규

1.문제

처음 파일을 받아보면, JPEG를 알고 있는지에 대한 질문과는 다르게 EXE 파일이 주어진다.



실행하면, `JPEG의 시그니처`를 알고 있는지에 대한 Alert가 발생하고, 확인을 누르면 종료된다.



2. 풀이

2.1 JPEG 시그니처

파일은 자신을 어떻게 인식시켜야 하는지에 대한 파일 타입 명세서를 가지고 있다. 이는 확장자로 표현되기도 하지만, 가장 정확한 표현은 `파일 헤더`이다. 예를 들어, 파일을 HEX로 보았을 때, 처음 나타나는 BYTE가 해당 파일의 타입을 나타낸다.

여기서 JPEG를 나타내는 시그니처는 `FF D8 FF E0` 이다.

먼저 Hxd같은 툴을 이용하여 `JPEG.exe`를 열어보면 다음과 같은 데이터들이 존재한다.

```
Offset(h)  00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F  Decoded text
00000000   4D 5A 90 00 03 00 00 00 04 00 00 00 FF FF 00 00  MZ..........ÿÿ..
00000010   B8 00 00 00 00 00 00 00 40 00 00 00 00 00 00 00  ¸.......@.......
00000020   00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00  ................
00000030   00 00 00 00 00 00 00 00 00 00 00 00 80 00 00 00  ............€...
00000040   0E 1F BA 0E 00 B4 09 CD 21 B8 01 4C CD 21 54 68  ..º..´.Í!¸.LÍ!Th
00000050   69 73 20 70 72 6F 67 72 61 6D 20 63 61 6E 6E 6F  is program canno
00000060   74 20 62 65 20 72 75 6E 20 69 6E 20 44 4F 53 20  t be run in DOS
00000070   6D 6F 64 65 2E 0D 0D 0A 24 00 00 00 00 00 00 00  mode....$.......
00000080   50 45 00 00 64 86 11 00 4F F0 04 66 00 80 01 00  PE..d†..Oð.f.€..
00000090   D6 05 00 00 F0 00 27 00 0B 02 02 18 00 1E 00 00  Ö...ð.'.........
```

- 하지만 JPEG의 시그니처가 보이지 않는다.

JPEG시그니처를 검색하기 위해 `FF D8 FF E0`을 HEX값으로 조회해 본다.

```
000201D0   65 66 70 74 72 2E 5F 5F 78 63 5F 7A 00 5F 5F 5F  efptr.__xc_z.___
000201E0   63 72 74 5F 78 74 5F 65 6E 64 5F 5F 00 5F 5F 6C  crt_xt_end__.__l
000201F0   69 62 36 34 5F 6C 69 62 6D 73 76 63 72 74 5F 61  ib64_libmsvcrt_a
00020200   5F 69 6E 61 6D 65 00 5F 5F 73 65 63 75 72 69 74  _iname.__securit
00020210   79 5F 63 6F 6F 6B 69 65 00 00 00 00 00 00 00 00  y_cookie........
00020220   FF D8 FF E0 00 10 4A 46 49 46 00 01 01 00 00 00  ÿØÿà..JFIF......
00020230   00 00 00 00 FF E1 00 62 45 78 69 66 00 00 4D 4D  ....ÿá.bExif..MM
00020240   00 2A 00 00 00 08 00 05 01 12 00 03 00 00 00 01  .*..............
00020250   00 01 00 00 01 1A 00 05 00 00 00 01 00 00 00 4A  ...............J
00020260   01 1B 00 05 00 00 00 01 00 00 00 52 01 28 00 03  ...........R.(..
00020270   00 00 00 01 00 01 00 00 02 13 00 03 00 00 00 01  ................
```

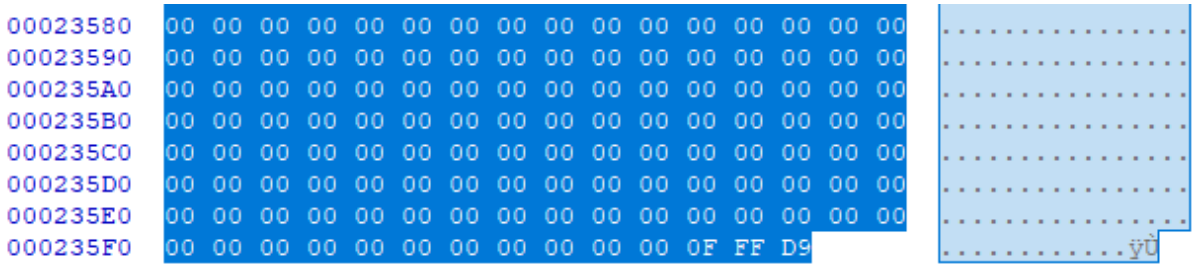
- offset `0x20220`에 JPEG 헤더가 발견되었다!

2.2 JPEG 추출

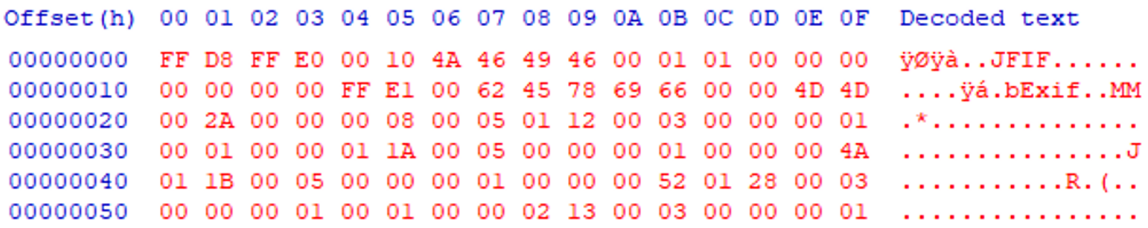
추출하는 방법은 두 가지가 있다. 첫번째는 시그니처 부분부터 데이터를 긁어 새로운 파일을 만드는 것이고, 두번째는 binwalk같은 파일 카빙 도구를 사용하여 추출하는 것이다.

Manual하게 데이터 추출

JPEG 데이터의 끝 부분을 나타내는 `FF D9` 까지 데이터를 전부 긁습니다.

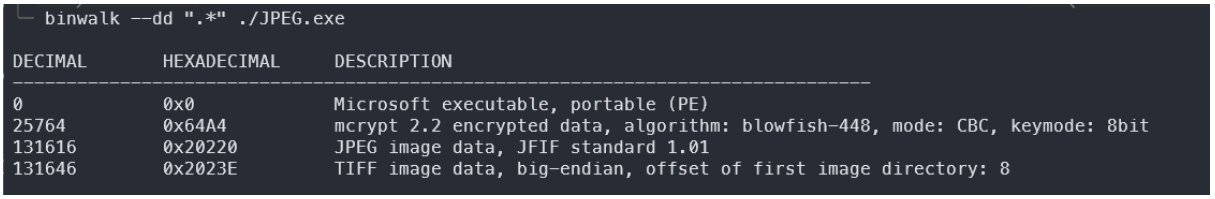


새로운 파일을 생성하고, 헤더부터 푸터까지 긁은 데이터를 복사 붙여 넣기 하면 됩니다.



binwalk를 이용한 추출

`binwalk --dd ".*" ./파일이름` 명령어를 통해, 모든 시그니처를 추출합니다.



- JPEG 파일이 있다고 하네요!

추출 이후에는, HEXADECIMAL에 대한 이름으로 파일이 생성됩니다.

- `0X20220` 형식으로 JPEG가 카빙되었으니, 확장자를 바꿔보겠습니다.

2.3 결과

manual 하게 추출하거나 binwalk를 사용하여 추출할 경우, 다음과 같은 JPEG 파일을 확인할 수 있습니다.