안녕하십니까. 악성코드 분석 스터디를 하고 있는 23학번 김동환입니다. 4학년 김동건 선배랑 두명이서 스터디 하고 있고, 지금까지 스터디 한 내용에 대해 발표하고자 합니다.

발표 내용의 Index로는 처음으론 PE 구조가 무엇인지, PE구조 중 중요한 부분에 대해 간략히 설명, 그리고 이를 출력하기 위한 코드 짠 것을 보여드리는 순이 되겠습니다.

PE 파일 포맷이란 윈도우에서 exe같은 실행계열이나 라이브러리 계열 등의 파일을 위한 파일 형식입니다.

구조는 직관적으로 표시해보면 사이즈가 좀 작지만 ppt의 왼쪽과 같은 그림으로 표현될 수 있고, 그 구성요소는 ms dos header, pe header, section header, section 이렇게 크게 4가지로 구분 가능하며 전부 16진수로 이루어져 있습니다.

MS DOS HEADER는 윈도우 초기에 주로 사용되던 것으로 범용성을 위해 남겨져있으며 주로 DOS환경에서 사용이 됩니다.

왼쪽위에 빨간 부분은 이의 magic code로 big endian으로 아스키코드 변환 시 MZ라고 나옵니다.

오른쪽 하단은 e\_lfanew라는, ms dos header 다음에 있는 pe header의 시작 부분 file offset의 값을 갖고 있으며, 윈도우 11일 경우에는 16진수 F0의 값을 가집니다.

바로 다음엔 ms dos stub이 이어나오는데 그렇게 중요한 부분은 아니니 넘어가겠습니다.

PE 헤더는 signature 구문, 그리고 pe file header, pe optional header란 2개의 구조체로 크게 3가지로 볼 수 있으며, signature은 PE . . 으로 나타납니다.

그리고 지금부터 설명할 때는 최대한 간추려 말씀드릴건데, Pe file header에서 machine은 프로그램이 32 or 64, number of section은 pe header 다음에 section header 다음에 오는 section의 개수. Size of optional header는 pe optional header의 16진수 전체 사이즈 총 240바이트구요, characteristic은 pe file header의 특징입니다.

Pe optional header는 magic code, address of entry point는 최초 실행되는 코드의 시작주소값. Imagebase는 가상메모리 내에서 pe file이 매핑되는 시작위치, 파일에서 section의 최소단위. Number of rva and sizes는 pe optional header의 뒷부분인 image data directory의 개수.라고 뭐라뭐라 설명.

그리고 image data directory는 보시다시피 typedef하고 data directory [16]으로 구조체 내에 구조체배열이 총 16개 있는 상태입니다. 대부분 DLL등을 포함하는 데 사용됩니다. 근데 IAT에 대해서는 아직 제대로 공부를 안해서 일단 이렇게 넘어가겠습니다.

그리고 Section header같은 경우에는 SECTION의 메타적인 정보를 저장하는 구간으로 Virtual Size는 VA에서 Section의 크기, Virtual Address는 말 그대로 VA, 즉 VA의 위치, Size of RAW DATA는 File에서의 section의 크기, pointer to raw data는 File에서의 section이 시작하는 flie offset characteristic은 각 section의 특징. 정도로 정리가능합니다.

그리고 section같은 경우는 예시를 하나만 드는 게 더 나을 거 같아서. 그림의 text section header를 보면 pointer to raw data(보라색)이 0x1000이고, size of raw data가 0x28000이라