



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0104047
(43) 공개일자 2024년07월04일

- | | |
|--|---|
| <p>(51) 국제특허분류(Int. Cl.)
G06F 18/2413 (2023.01) G06F 18/20 (2023.01)
G06F 18/23 (2023.01) G06F 21/56 (2013.01)
G06N 20/20 (2019.01)</p> <p>(52) CPC특허분류
G06F 18/2413 (2023.01)
G06F 18/23 (2023.01)</p> <p>(21) 출원번호 10-2023-0191930
(22) 출원일자 2023년12월26일
심사청구일자 2023년12월26일</p> <p>(30) 우선권주장
1020220185592 2022년12월27일 대한민국(KR)</p> | <p>(71) 출원인
충남대학교산학협력단
대전광역시 유성구 대학로 99 (궁동, 충남대학교)</p> <p>(72) 발명자
김형식
대전광역시 유성구 대학로 99 충남대학교
조우진
대전광역시 유성구 대학로 99 충남대학교</p> <p>(74) 대리인
제일특허법인(유)</p> |
|--|---|

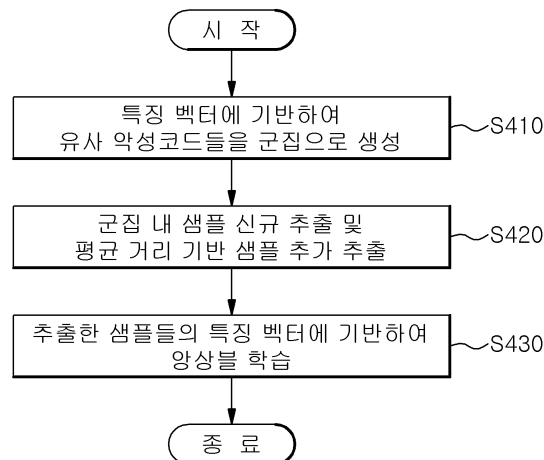
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 악성코드 탐지 장치 및 그 학습 방법

(57) 요약

일 실시예에 따른 악성코드 탐지 장치의 학습 방법은, 학습셋으로부터 특징 벡터에 기반하여 유사 악성코드들을 그룹화하여 군집으로 생성하는 단계; 상기 군집 내에서 무작위로 샘플을 신규 추출한 후 기 추출된 샘플과의 평균 거리에 기반하여 샘플을 추가 추출하는 단계; 및 상기 신규 추출 및 상기 추가 추출한 샘플들의 상기 특징 벡터에 기반하여 악성코드 탐지 모델을 앙상블 학습하는 단계;를 포함한다.

대표도 - 도4



(52) CPC특허분류

G06F 18/285 (2023.01)

G06F 21/56 (2013.01)

G06N 20/20 (2021.08)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711160654
과제번호	2022-0-01200-001
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	융합보안대학원(충남대학교)
기 여 율	1/1
과제수행기관명	충남대학교 산학협력단
연구기간	2022.01.01 ~ 2022.12.31

명세서

청구범위

청구항 1

악성코드 탐지 장치의 학습 방법으로서,

학습셋으로부터 특징 벡터에 기반하여 유사 악성코드들을 그룹화하여 군집으로 생성하는 단계;

상기 군집 내에서 샘플을 신규 추출 및 추가 추출하는 단계; 및

상기 신규 추출 및 상기 추가 추출한 샘플들의 상기 특징 벡터에 기반하여 악성코드 탐지 모델을 앙상블 학습하는 단계;를 포함하는

학습 방법.

청구항 2

제 1항에 있어서,

상기 생성하는 단계에서 상기 그룹화의 결과에 따라 군집별 군집번호 및 군집에 포함된 샘플 번호가 등재된 샘플 번호 디렉터리를 생성하는

학습 방법.

청구항 3

제 2 항에 있어서,

상기 추출하는 단계는, 상기 군집번호를 대상으로 복원 추출을 수행하여 중복을 포함한 군집 번호 리스트를 생성하며, 상기 군집 번호 리스트의 군집 번호를 기반으로 상기 샘플을 추출하여 군집 번호별 샘플 번호 리스트를 생성해 상기 군집 번호별 샘플 번호 리스트가 상기 앙상블 학습에 이용되도록 하는

학습 방법.

청구항 4

제 3 항에 있어서,

상기 군집 번호 리스트를 대상으로 복원 추출을 군집 번호 개수만큼 수행하여 상기 중복을 포함하는 군집 번호 리스트를 생성하는

학습 방법.

청구항 5

제 3 항에 있어서,

상기 샘플을 신규 추출시에 무작위로 추출하고, 상기 샘플을 추가 추출 시 기 추출된 샘플과의 평균 거리가 가장 먼 샘플을 상기 특징 벡터에 기반하여 추출하는

학습 방법.

청구항 6

제 1 항에 있어서,

상기 앙상블 학습하는 단계 이후, 상기 악성코드 탐지 모델 중 특정 샘플에 과적합 된 모델을 제거하는 단계를 더 포함하는

학습 방법.

청구항 7

제 6 항에 있어서,

상기 제거하는 단계는, 상기 악성코드 탐지 모델을 대상으로 가장 높은 중요도를 가지는 패턴을 추출하여 해당 패턴이 학습셋에 포함되지 않은 샘플들로 구성된 검증셋에 조회했을 때 존재하지 않으면 제거하는

학습 방법.

청구항 8

적어도 하나의 명령어를 저장하는 메모리; 및

프로세서;를 포함하며,

상기 프로세서는, 상기 적어도 하나의 명령어를 실행함으로써, 기 학습된 악성코드 탐지 모델이 악성코드 탐지를 수행하고,

상기 악성코드 탐지 모델은,

학습셋으로부터 특징 벡터에 기반하여 유사 악성코드들이 그룹화되어 군집으로 생성되며, 상기 군집 내에서 샘플이 신규 추출 및 추가 추출되고, 상기 신규 추출 및 상기 추가 추출된 샘플들의 상기 특징 벡터에 기반하여 상기 악성코드 탐지 모델이 앙상블 학습된

악성 코드 탐지 장치.

청구항 9

컴퓨터 판독 가능 기록매체에 저장되어 있는 컴퓨터 프로그램으로서,

상기 컴퓨터 프로그램은, 프로세서에 의해 실행되면,

학습셋으로부터 특징 벡터에 기반하여 유사 악성코드들을 그룹화하여 군집으로 생성하는 단계;

상기 군집 내에서 샘플을 신규 추출 및 추가 추출하는 단계; 및

상기 신규 추출 및 상기 추가 추출한 샘플들의 상기 특징 벡터에 기반하여 악성코드 탐지 모델을 앙상블 학습하는 단계;를 포함하는 악성코드 탐지 장치의 학습 방법을 상기 프로세서가 수행하도록 하기 위한 명령어를 포함하는

컴퓨터 판독 가능 기록매체에 저장되어 있는 컴퓨터 프로그램.

청구항 10

컴퓨터 프로그램이 저장되어 있는 컴퓨터 판독 가능 기록매체로서,

상기 컴퓨터 프로그램은, 프로세서에 의해 실행되면,

학습셋으로부터 특징 벡터에 기반하여 유사 악성코드들을 그룹화하여 군집으로 생성하는 단계;

상기 군집 내에서 샘플을 신규 추출 및 추가 추출하는 단계; 및

상기 신규 추출 및 상기 추가 추출한 샘플들의 상기 특징 벡터에 기반하여 악성코드 탐지 모델을 앙상블 학습하는 단계;를 포함하는 악성코드 탐지 장치의 학습 방법을 상기 프로세서가 수행하도록 하기 위한 명령어를 포함하는

컴퓨터 프로그램이 저장되어 있는 컴퓨터 판독 가능 기록매체.

발명의 설명**기술 분야**

본 발명은 악성코드 탐지를 위한 학습 방법과 이 방법으로 학습된 악성코드 탐지 장치에 관한 것이다.

[0001]

배경 기술

- [0002] 급증하는 악성코드에 대응하기 위해 기계학습에 기반한 탐지 기법이 연구 및 개발됐다. 기계학습 기법의 하나인 분류 기법은 악성코드와 정상코드에서 나타나는 패턴을 자동으로 학습할 수 있는 기법으로, 학습된 모델을 이용하여 악성코드를 탐지하는 것이 가능하다. 그렇지만, 모델을 학습하기 위해서는 악성코드와 정상코드를 포함한 학습셋이 필요한데, 대부분의 악성코드는 기존의 악성코드에서 파생된 유사 악성코드이기에 정상적인 패턴을 학습하기 어렵다는 문제가 있다. 왜냐하면, 학습 알고리즘은 반복되는 패턴을 학습하기에 중복되는 특성을 공유하는 유사 악성코드들에 편향될 수 있기 때문이다. 즉, 유사 악성코드가 포함된 학습셋을 이용해 학습할 경우 유사 악성코드들에서 반복되는 의미 없는 패턴을 학습하여 일반화 성능이 떨어지는 문제점이 있다.
- [0003] 이러한 문제점을 해소하기 위해서는 학습셋에 포함된 유사 악성코드들을 제거해야 하지만, 유사한 기준을 정확히 정의할 수 없기 때문에 유사 악성코드들이 포함된 학습셋을 이용해 일반화된 모델을 학습할 수 있는 방법이 필요한 실정이다.

선행기술문헌

특허문헌

- [0004] (특허문헌 0001) 공개특허공보 제10-2022-0060203호, 공개일 2022년 05월 11일.

발명의 내용

해결하려는 과제

- [0005] 유사 악성코드가 포함된 학습셋을 이용해 일반화된 모델을 학습하기 위한 방안으로 학습셋으로부터 샘플 추출을 통해 새로운 학습셋을 다수 생성하고, 생성된 학습셋들로부터 모델을 학습하여 종합하는 앙상블 학습을 고려할 수 있다. 그렇지만, 악성코드가 포함된 학습셋의 경우 유사 악성코드에 편향되었기 때문에 단순히 무작위 추출을 수행할 경우 여전히 유사 악성코드들이 편향된 학습셋이 생성된다. 왜냐하면 무작위 추출의 경우 모집단의 분포를 그대로 유지하는 특성이 있는데, 학습셋에 특정 유사 악성코드의 집합이 큰 비중을 차지할 경우 무작위 추출을 수행하더라도 여전히 해당 집합이 큰 비중을 차지하기 때문이다. 이에 유사 악성코드가 포함된 학습셋을 대상으로 샘플 추출을 수행할 때 유사 악성코드를 최대한 배제하고 다양한 악성코드로 구성된 새로운 학습셋을 생성하여 앙상블 학습을 수행하는 방법이 필요한 실정이다.
- [0006] 일 실시예에 따르면, 학습셋에 포함된 유사 악성코드들을 그룹화하여 군집 단위로 샘플을 추출하여 같은 군집 내의 유사 악성코드가 추출될 가능성을 최대한 배제하고, 같은 군집 내에서도 이미 추출된 샘플들과 평균 거리가 가장 먼 샘플을 추출하여 다양한 악성코드들을 추출해 모델을 학습하는 악성코드 탐지 장치 및 그 학습 방법을 제공한다.
- [0007] 본 발명의 해결하고자 하는 과제는 이상에서 언급한 것으로 제한되지 않으며, 언급되지 않은 또 다른 해결하고자 하는 과제는 아래의 기재로부터 본 발명이 속하는 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0008] 제 1 관점에 따른 악성코드 탐지 장치의 학습 방법은, 학습셋으로부터 특징 벡터에 기반하여 유사 악성코드들을 그룹화하여 군집으로 생성하는 단계; 상기 군집 내에서 무작위로 샘플을 신규 추출한 후 기 추출된 샘플과의 평균 거리에 기반하여 샘플을 추가 추출하는 단계; 및 상기 신규 추출 및 상기 추가 추출된 샘플들의 상기 특징 벡터에 기반하여 악성코드 탐지 모델을 앙상블 학습하는 단계;를 포함한다.
- [0009] 제 2 관점에 따른 악성코드 탐지 장치는, 적어도 하나의 명령어를 저장하는 메모리; 및 프로세서;를 포함하며, 상기 프로세서는, 상기 적어도 하나의 명령어를 실행함으로써, 기 학습된 악성코드 탐지 모델이 악성코드 탐지를 수행하고, 상기 악성코드 탐지 모델은, 학습셋으로부터 특징 벡터에 기반하여 유사 악성코드들이 그룹화되어 군집으로 생성되며, 상기 군집 내에서 무작위로 샘플이 신규 추출된 후 기 추출된 샘플과의 평균 거리에 기반하여 샘플이 추가 추출되고, 상기 신규 추출 및 상기 추가 추출된 샘플들의 상기 특징 벡터에 기반하여 상기 악성

코드 탐지 모델이 앙상블 학습된다.

[0010] 제 3 관점에 따른 컴퓨터 판독 가능 기록매체에 저장되어 있는 컴퓨터 프로그램의 상기 컴퓨터 프로그램은, 프로세서에 의해 실행되면, 상기 악성코드 탐지 장치의 학습 방법을 상기 프로세서가 수행하도록 하기 위한 명령어를 포함한다.

[0011] 제 4 관점에 따른 컴퓨터 프로그램이 저장되어 있는 컴퓨터 판독 가능 기록매체의 상기 컴퓨터 프로그램은, 프로세서에 의해 실행되면, 상기 악성코드 탐지 장치의 학습 방법을 상기 프로세서가 수행하도록 하기 위한 명령어를 포함한다.

발명의 효과

[0012] 일 실시예에 의하면, 학습셋에 포함된 유사 악성코드들을 그룹화하여 군집 단위로 샘플을 추출하여 같은 군집 내의 유사 악성코드가 추출될 가능성을 최대한 배제하고, 같은 군집 내에서도 이미 추출된 샘플들과 평균 거리가 가장 먼 샘플을 추출함으로써, 다양한 악성코드들을 추출해 악성코드 탐지 모델을 학습하여 높은 악성코드 탐지 성능을 제공할 수 있는 효과가 있다.

도면의 간단한 설명

[0013] 도 1은 본 발명의 일 실시예에 따른 악성코드 탐지를 위한 학습 방법으로 학습될 수 있는 악성코드 탐지 장치의 구성도이다.

도 2는 도 1에 도시된 악성코드 탐지 장치 내 악성코드 탐지 모델의 기능을 나타내는 구성도이다.

도 3은 쿠크 샌드박스 보고서를 파싱(parsing)한 예시도이다.

도 4는 본 발명의 일 실시예에 따른 악성코드 탐지를 위한 학습 방법을 설명하기 위한 흐름도이다.

도 5는 도 2에 나타난 악성코드 탐지 장치 내 악성코드 탐지 모델이 악성코드 탐지를 위한 학습 방법을 수행하는 과정을 예시한 개념도이다.

발명을 실시하기 위한 구체적인 내용

[0014] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다.

[0015] 본 명세서에서 사용되는 용어에 대해 간략히 설명하고, 본 발명에 대해 구체적으로 설명하기로 한다.

[0016] 본 발명에서 사용되는 용어는 본 발명에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어들을 선택하였으나, 이는 당 분야에 종사하는 기술자의 의도 또는 관례, 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 특정한 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 발명의 설명 부분에서 상세히 그 의미를 기재할 것이다. 따라서 본 발명에서 사용되는 용어는 단순한 용어의 명칭이 아닌, 그 용어가 가지는 의미와 본 발명의 전반에 걸친 내용을 토대로 정의되어야 한다.

[0017] 명세서 전체에서 어떤 부분이 어떤 구성요소를 ‘포함’ 한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있음을 의미한다.

[0018] 또한, 명세서에서 사용되는 ‘부’라는 용어는 소프트웨어 또는 FPGA나 ASIC과 같은 하드웨어 구성요소를 의미하며, ‘부’는 어떤 역할들을 수행한다. 그렇지만 ‘부’는 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. ‘부’는 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 ‘부’는 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로 코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들 및 변수들을 포함한다. 구성요소들과 ‘부’들 안에서 제공되는 기능은 더 작은 수의 구성요소들 및 ‘부’들로 결합되거나 추가적인 구성요소들과 ‘부’들로 더 분리될 수 있다.

[0019] 아래에서는 첨부한 도면을 참고하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을

가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략한다.

- [0020] 도 1은 본 발명의 일 실시예에 따른 악성코드 탐지를 위한 학습 방법으로 학습될 수 있는 악성코드 탐지 장치의 구성도이다.
- [0021] 도 1을 참조하면, 악성코드 탐지 장치(100)는 악성코드 탐지 모델(111)이 탑재된 메모리(110) 및 프로세서(120)를 포함하고, 입력부(130) 및/또는 출력부(140)를 더 포함할 수 있다.
- [0022] 메모리(110)에는 프로세서(120)가 실행할 수 있는 명령어를 포함하는 악성코드 탐지 모델(111)이 탑재되고, 악성코드 탐지 모델(111)의 실행에 한 각종 정보를 더 저장할 수 있다. 이러한 악성코드 탐지 모델(111)은 프로세서(120)에 의해 실행됨으로써, 프로세서(120)가 학습셋으로부터 특징 벡터에 기반하여 유사 악성코드들을 군집으로 생성하고, 군집 내에서 무작위로 샘플을 신규 추출한 후 기 추출된 샘플과의 평균 거리에 기반하여 샘플을 추가 추출하며, 신규 추출 및 추가 추출한 샘플들의 특징 벡터에 기반하여 악성코드 탐지 모델을 앙상블 학습하도록 할 수 있다.
- [0023] 프로세서(120)는 메모리(110)에 저장된 명령어를 실행함으로써 악성코드 탐지 모델(111)을 로드하여, 악성코드 탐지를 위한 학습과 학습 후 악성코드 탐지를 수행할 수 있다. 이러한 프로세서(120)는 하나 또는 복수의 프로세서로 구성될 수 있다. 예를 들어, 하나 또는 복수의 프로세서는 CPU(central process unit), DSP(digital signal processor) 등과 같은 범용 프로세서, GPU(graphic process unit), VPU(vision process unit) 등과 같은 그래픽 전용 프로세서 또는 NPU(neural process unit)와 같은 인공지능 전용 프로세서일 수 있다.
- [0024] 프로세서(120)는 메모리(110)에 탑재된 악성코드 탐지 모델(111)을 로드해 실행함으로써, 학습셋으로부터 특징 벡터에 기반하여 유사 악성코드들을 군집으로 생성하고, 군집 내에서 무작위로 샘플을 신규 추출한 후 기 추출된 샘플과의 평균 거리에 기반하여 샘플을 추가 추출하며, 신규 추출 및 추가 추출한 샘플들의 특징 벡터에 기반하여 악성코드 탐지 모델을 앙상블 학습한다. 여기서, 프로세서(120)는 주어진 프로그램을 분석한 결과에 따라 특징 벡터를 정의할 수 있고, 정의된 특징 벡터에 기반하여 유사 악성코드들을 하나의 군집으로 그룹화할 수 있다. 또한, 프로세서(120)는 특징 벡터를 각 프로그램을 대상으로 특징 리스트에 정의된 각 특징의 발생 빈도를 측정하여 생성할 수 있다. 예를 들어, 프로세서(120)는 그룹화의 결과에 따라 군집별 군집번호 및 군집에 포함된 샘플 번호가 등재된 샘플 번호 디렉토리를 생성할 수 있고, 군집번호를 대상으로 복원 추출을 수행하여 중복을 포함한 군집 번호 리스트를 생성할 수 있으며, 군집 번호 리스트의 군집 번호를 기반으로 샘플을 추출하여 군집 번호별 샘플 번호 리스트를 생성해 군집 번호별 샘플 번호 리스트가 앙상블 학습에 이용되도록 할 수 있다. 예컨대, 프로세서(120)는 군집 번호 리스트를 대상으로 복원 추출을 군집 번호 개수만큼 수행하여 중복을 포함하는 군집 번호 리스트를 생성할 수 있고, 샘플을 추가 추출할 때에 기 추출된 샘플과의 평균 거리가 가장 먼 샘플을 특징 벡터에 기반하여 추출할 수 있다.
- [0025] 또한, 프로세서(120)는 앙상블 학습 이후, 악성코드 탐지 모델 중 특정 샘플에 과적합 된 모델을 제거할 수 있다. 예를 들어, 프로세서(120)는 악성코드 탐지 모델을 대상으로 가장 높은 중요도를 가지는 패턴을 추출하여 해당 패턴이 학습셋에 포함되지 않은 샘플들로 구성된 검증셋(validation set)에 조회했을 때 존재하지 않으면 제거할 수 있다.
- [0026] 입력부(130)는 프로세서(120)가 악성코드 탐지 모델(111)을 실행하는데 필요한 각종 정보를 입력 받을 수 있다. 이러한 각종 정보는 실시간 입력될 수 있고, 미리 입력된 경우는 각종 정보가 메모리(110)에 저장될 수 있다. 예를 들어, 입력부(130)는 학습셋이나 악성코드 탐지 대상 프로그램 등을 입력받을 수 있다.
- [0027] 출력부(140)는 프로세서(120)가 악성코드 탐지 모델(111)을 실행함에 따라 발생하는 각종 처리 데이터를 출력할 수 있다. 예를 들어, 출력부(140)는 각종 처리 데이터를 외부 주변기기 등으로 출력할 수 있는 데이터 인터페이스, 각종 처리 데이터를 통신 채널을 통해 송신할 수 있는 통신모듈 등을 포함할 수 있다.
- [0028] 도 2는 도 1에 도시된 악성코드 탐지 장치(100) 내 악성코드 탐지 모델(111)의 기능을 나타내는 구성도이다.
- [0029] 도 2를 참조하면, 악성코드 탐지 모델(111)은 특징 정의부(210), 군집 생성부(220), 군집 추출부(230), 샘플 추출부(240) 및 모델 관리부(250)를 포함할 수 있다. 이처럼, 도 2에 도시한 특징 정의부(210), 군집 생성부(220), 군집 추출부(230), 샘플 추출부(240) 및 모델 관리부(250)는 악성코드 탐지 모델(111)에 대한 이해를 돕기 위해 악성코드 탐지 모델(111)의 기능을 나눈 것으로서, 이에 한정되지 않는다. 실시예들에 따라, 특징 정의부(210), 군집 생성부(220), 군집 추출부(230), 샘플 추출부(240) 및 모델 관리부(250)에 의한 각각의 기능은

병합 가능하며, 하나의 프로그램에 포함된 일련의 명령어들로 구현될 수 있다.

- [0030] 악성코드 탐지 모델(111)의 특징 정의부(210)는 주어진 프로그램을 분석한 결과에 따라 특징 벡터를 정의한다. 예를 들어, 특징 벡터는 각 프로그램을 대상으로 특징 리스트에 정의된 각 특징의 발생 빈도를 측정하여 생성할 수 있다.
- [0031] 악성코드 탐지 모델(111)의 군집 생성부(220)는 학습셋으로부터 특징 벡터에 기반하여 유사 악성코드들을 그룹화하여 군집으로 생성한다. 예를 들어, 군집 생성부(220)는 그룹화의 결과에 따라 군집별 군집번호 및 군집에 포함된 샘플 번호가 등재된 샘플 번호 디서너리를 생성할 수 있다.
- [0032] 악성코드 탐지 모델(111)의 군집 추출부(230)는 샘플 번호 디서너리의 군집번호를 대상으로 복원 추출을 수행하여 중복을 포함한 군집 번호 리스트를 생성할 수 있다. 예를 들어, 군집 추출부(230)는 군집 번호 리스트를 대상으로 복원 추출을 군집 번호 개수만큼 수행하여 중복을 포함하는 군집 번호 리스트를 생성할 수 있다.
- [0033] 악성코드 탐지 모델(111)의 샘플 추출부(240)는 군집 내에서 무작위로 샘플을 신규 추출한 후 기 추출된 샘플과의 평균 거리에 기반하여 샘플을 추가 추출할 수 있다. 예를 들어, 샘플 추출부(240)는 군집 번호 리스트의 군집 번호를 기반으로 군집에 포함되는 샘플을 추출하는데, 군집으로부터 최초로 샘플이 추출되는 경우 무작위로 추출하지만, 이후 기 추출된 샘플이 있을 경우 기 추출된 샘플들과의 평균 거리가 가장 먼 샘플을 특징 벡터에 기반하여 추출할 수 있다. 이때, 기 추출된 샘플들을 관리하기 위해 군집 번호별 샘플 번호 리스트를 생성해 샘플이 추출될 때마다 해당되는 군집 번호의 샘플 번호 리스트에 추가하여 군집 번호별 샘플 번호 리스트가 이후 앙상블 학습에 이용되도록 할 수 있다.
- [0034] 악성코드 탐지 모델(111)의 모델 관리부(250)는 샘플 추출부(240)가 신규 추출 및 추가 추출한 샘플들의 특징 벡터에 기반하여 악성코드 탐지 모델(111)을 앙상블 학습할 수 있다. 또한, 모델 관리부(250)는 앙상블 학습한 이후, 악성코드 탐지 모델(111) 중 특정 샘플에 과적합 된 모델을 제거할 수 있다. 예를 들어, 모델 관리부(250)는 악성코드 탐지 모델(111)에 포함된 모델들을 대상으로 가장 높은 중요도를 가지는 패턴을 추출하여 해당 패턴이 학습셋에 포함되지 않은 샘플들로 구성된 검증셋(201)에 조회했을 때 존재하지 않으면 제거할 수 있다.
- [0035] 도 3은 쿠크 샌드박스 보고서를 파싱(parsing)한 예시도이고, 도 4는 본 발명의 일 실시예에 따른 악성코드 탐지를 위한 학습 방법을 설명하기 위한 흐름도이며, 도 5는 도 2에 나타난 악성코드 탐지 장치 내 악성코드 탐지 모델이 악성코드 탐지를 위한 학습 방법을 수행하는 과정을 예시한 개념도이다.
- [0036] 이하, 도 1 내지 도 5를 참조하여 본 발명의 일 실시예에 따른 악성코드 탐지 장치가 악성코드 탐지를 위해 학습을 수행하는 과정에 대해 자세히 살펴보기로 한다.
- [0037] 먼저, 악성코드 탐지 장치(100)의 프로세서(120)는 메모리(110)에 저장된 명령어를 실행함으로써 악성코드 탐지 모델(111)을 로드하여, 악성코드 탐지를 위한 학습을 수행할 수 있는 환경에 놓인다.
- [0038] 악성코드 탐지 모델(111)의 학습 과정을 살펴보면, 악성코드 탐지 모델(111)의 특징 정의부(210)는 주어진 프로그램을 분석한 결과에 따라 특징 벡터를 정의한다. 예를 들어, 특징 벡터는 각 프로그램을 대상으로 특징 리스트에 정의된 각 특징의 발생 빈도를 측정하여 생성할 수 있다.
- [0039] 특징 정의부(210)는 학습셋에 포함된 프로그램으로부터 데이터를 추출하고, 추출된 데이터를 조합한 특징을 정의하여 특징 리스트를 생성하고, 각 프로그램에 대하여 특징 벡터(V)를 생성할 수 있다. 특징 리스트는 학습셋에서 나타나는 모든 데이터를 합집합 함으로써 정의할 수 있으며, 특징 벡터(V)는 각 프로그램을 대상으로 특징 리스트에 정의된 각 특징의 발생 빈도를 측정하여 생성할 수 있다.
- [0040] 일 실시예에 따르면, 프로그램들이 실행되는 동안 호출하는 API들을 모두 합집합하여 특징 리스트를 정의할 수 있으며, 특징 벡터(V)는 각 프로그램에 대하여 특징 리스트에 정의된 API의 발생 빈도를 측정하여 생성할 수 있다. 프로그램이 실행되는 동안 호출하는 API를 수집하기 위해서는 악성코드를 실제로 실행시켜 행위를 감시할 필요가 있는데, 이러한 도구 중 하나로 쿠크 샌드박스(cuckoo sandbox)를 사용할 수 있다. 쿠크 샌드박스는 악성코드를 외부 환경과 격리된 가상머신에서 실행시켜 json 형태의 보고서를 결과로 낸다. 도 3은 쿠크 샌드박스 보고서를 파싱(parsing)한 예시이며, “JSON/behavior/processes/[1]/calls” 경로에서 특징값을 정의하기 위한 API 명칭, API에 정의된 인자, 전달되는 매개변수를 수집할 수 있다.
- [0041] 일 실시예에 따르면, 학습셋에 포함된 6319개의 프로그램을 쿠크 샌드박스를 통해 실행시켰을 때 호출된 API들을 합집합 하면 길이가 309인 특징 리스트가 생성된다. 그리고 각 프로그램에 대하여 특징 리스트에 정의된 각

특징의 발생 빈도를 측정하면 특징 벡터(V)는 표 1과 같이 표현될 수 있다. 표 1은 총 311개의 열로 구성되었는데, 1열은 각 프로그램의 식별번호이며, 마지막 열은 해당 프로그램의 정상과 악성 여부를 알려주는 라벨이다 (0: 정상, 1: 악성). 그리고, 2번 열부터 310번 열까지 총 309개의 열은 전체 특징 집합을 의미한다.

표 1

Id	NtQueryAttributesFile	DnsQuery_W	CopyFileW	...	LoadStringW	label
1	6	0	65		0	1
2	0	0	8		1	0
3	5	5	6		1	0
...						
6318	4	0	8		31	1
6319	4	0	0		7	1

군집 생성부(220)는 학습셋으로부터 특징 벡터에 기반하여 유사 악성코드들을 그룹화하여 군집으로 생성한다. 예를 들어, 군집 생성부(220)는 그룹화의 결과에 따라 군집별 군집번호 및 군집에 포함된 샘플 번호가 등재된 샘플 번호 디렉터리를 생성할 수 있다(S410).

군집 생성부(220)는 특징 정의부(210)에서 출력된 특징 벡터(V)에 기반하여 유사한 악성코드들을 군집화하고, 군집화된 결과를 가공하여 군집 별 포함된 샘플 번호 디렉터리(D1)를 출력할 수 있다. 여기서 군집화된 결과란 각 샘플에 해당하는 군집 번호를 차례로 저장한 군집 번호 리스트를 의미하는데, 같은 군집 번호를 가진 샘플은 같은 군집에 속한 것으로 간주한다. 이에 군집 번호 리스트로부터 같은 군집 번호를 가진 샘플들을 모으면 군집 별로 포함된 샘플 번호 디렉터리(D1)를 구축할 수 있다.

일 실시예에 따르면, 6319개의 특징 벡터(V)를 대상으로 계층적 클러스터링을 임계값 0.31로 수행할 경우 총 1149개의 군집이 생성되며, 각 샘플이 해당하는 군집 번호를 차례로 저장하면 길이가 6319인 군집 번호 리스트가 생성된다. 그리고, 같은 군집 번호가 할당된 샘플들을 모아 리스트로 나타내면 표 2와 같이 군집 번호를 Key로 갖고, 각 군집에 포함된 샘플 번호를 Value로 갖는 군집 별 포함된 샘플 디렉터리(D1)를 출력할 수 있다.

표 2

군집 번호	포함된 샘플 번호
1	[3029, 1333]
2	[5942, 233, 4223, 13, 59, 127, 1134]
3	[2888, 6008, 488, 987, 1422]
...	
49	[3998, 1774, 458]
1149	[377, 3322, 132]

악성코드 탐지 모델(111)의 군집 추출부(230)는 샘플 번호 디렉터리의 군집번호를 대상으로 복원 추출을 수행하여 중복을 포함한 군집 번호 리스트를 생성할 수 있다. 예를 들어, 군집 추출부(230)는 군집 번호 리스트를 대상으로 복원 추출을 군집 번호 개수만큼 수행하여 중복을 포함하는 군집 번호 리스트를 생성할 수 있다.

군집 추출부(230)는 군집 생성부(220)에서 출력된 군집 별 포함된 샘플 번호 디렉터리(D1)로부터 Key 값만 추출하여 군집 번호 리스트를 얻고, 군집 번호 리스트를 대상으로 복원 추출을 군집 번호 개수만큼 수행하여 중복을 포함한 군집 번호 리스트(L1)를 생성할 수 있다. 이때, 사용자가 설정한 횟수(N)만큼 이 과정을 반복하여 다수의 중복을 포함한 군집 번호 리스트(L1)를 출력할 수 있다.

일 실시예에 따르면, 군집 생성부(220)에서 출력된 표 2와 같은 군집 별 포함된 샘플 디렉터리(D1)로부터 Key 값만 추출하면 총 1149개의 군집 번호를 얻을 수 있으며, 1149개의 군집 번호를 대상으로 군집 번호 개수(1149개)만큼 복원 추출을 수행할 경우 길이가 1149인 중복을 포함한 군집 번호 리스트(L1)가 생성된다. 사용자가 반복 횟수(N)를 50으로 설정할 경우 표 3과 같이 50개의 리스트가 생성된다.

표 3

리스트 번호	중복을 포함한 군집 번호 리스트
1	[39, 39, 22, 229, ...]

2	[303, 734, 303, 18, ...]
3	[24, 145, 24, 677, ...]
...	
49	[14, 32, 14, 14, 14, ...]
50	[1142, 88, 1142, 1142, ...]

- [0051] 악성코드 탐지 모델(111)의 샘플 추출부(240)는 군집 내에서 무작위로 샘플을 신규 추출한 후 기 추출된 샘플과의 평균 거리에 기반하여 샘플을 추가 추출할 수 있다. 예를 들어, 샘플 추출부(240)는 군집 번호 리스트의 군집 번호를 기반으로 해당 군집에 포함된 샘플을 추출하여 군집 번호별 샘플 번호 리스트를 생성해 군집 번호별 샘플 번호 리스트가 이후 앙상블 학습에 이용되도록 할 수 있다. 또한, 샘플 추출부(240)는 샘플을 추가 추출할 때에 기 추출된 샘플과의 평균 거리가 가장 먼 샘플을 특징 벡터에 기반하여 추출할 수 있다(S420).
- [0052] 샘플 추출부(240)는 특징 정의부(210)에서 출력된 특징 벡터(V)와 군집 생성부(220)에서 출력된 군집 별 포함된 샘플 번호 디서너리(D1)와 군집 추출부(230)에서 출력된 다수의 중복을 포함한 군집 번호 리스트(L1)에 기반하여 최대한 다양한 샘플을 추출해 샘플 번호 리스트(L2)를 다수 출력할 수 있다. 결과적으로 다양한 샘플들을 이용해 일반화된 모델을 학습하여 앙상블 학습을 수행할 수 있다. 구체적으로는 군집 추출부(230)에서 출력된 중복을 포함한 군집 번호 리스트(L1)로부터 군집 번호를 차례로 선택하고, 선택된 군집 번호를 군집 생성부(220)에서 출력된 군집 별 포함된 샘플 번호 디서너리(D1)에 조회하여 해당 군집에 속하는 샘플 번호 리스트를 얻고, 선택된 군집에서 이미 추출된 샘플이 존재하지 않으면 무작위로 선택하지만, 존재할 경우 특징 정의부(210)에서 출력된 특징 벡터(V)에 기반하여 해당 샘플들과 평균 거리가 가장 먼 샘플을 추출한다. 만약, 해당 군집 번호에 속하는 샘플들이 모두 추출되었다면 이미 추출된 샘플 번호의 리스트를 초기화하고 무작위로 선택한다. 이때, 이미 추출된 샘플의 여부를 고려하기 위해 군집 별 추출된 샘플 번호 디서너리(D2)를 생성하여 각 군집 별 추출된 샘플들을 관리한다. 최종적으로 상기 군집 추출부에서 출력된 각 중복을 포함한 군집 번호 리스트(L1)별로 샘플 번호 리스트(L2)가 출력된다.
- [0053] 일 실시예에 따르면, 군집 추출부(230)에서 출력된 중복을 포함한 군집 번호 리스트(L2) 중 하나가 [14, 32, 14, 14, 14, 14] 와 같이 구성되고, 군집 생성부(220)에서 출력된 군집 별 샘플 번호 디서너리(D1)가 D1={14: [3488, 5322, 277, 1443], 32: [857, 6088, 4222]}와 같이 구성되고, 이미 추출된 샘플의 여부를 고려하기 위한 군집 별 추출된 샘플 번호 디서너리(D2)가 D2={14: [], 32: []}와 같이 구성된다고 할 때 L1에서 각 군집 번호가 선택될 때마다 수행되는 동작은 아래와 같다.
- [0054] 일 실시예에 따르면, 가장 먼저 L1에서 첫 번째 요소인 14가 선택되고 D1에서 14에 해당하는 [3488, 5322, 277, 1443] 중에서 하나를 선택하는데, D2에서 14에 해당하는 리스트는 비어있으므로 랜덤하게 선택된다. 만약 3488이 선택되었다고 하면 D2의 14에 해당하는 리스트에 3488을 삽입한다.
- [0055] 일 실시예에 따르면, L1에서 두 번째 요소인 32가 선택되고, D1에서 32에 해당하는 [857, 6088, 4222] 중에서 하나를 선택하는데, D2에서 32에 해당하는 리스트는 비어있으므로 랜덤하게 선택된다. 만약, 4222이 선택되었다고 하면 D2의 32에 해당하는 리스트에 4222를 삽입한다.
- [0056] 일 실시예에 따르면, L1에서 세 번째 요소인 14가 선택되고, D1에서 14에 해당하는 [3488, 5322, 277, 1443] 중에서 하나를 선택하는데, D2에서 14에 해당하는 리스트가 3488을 포함하고 있으므로, 3488과 가장 거리가 먼 샘플을 선택한다. 만약, 277이 선택되었다고 하면 D2의 14에 해당하는 리스트에 277을 삽입한다.
- [0057] 일 실시예에 따르면, L1에서 네 번째 요소인 14가 선택되고, D1에서 14에 해당하는 [3488, 5322, 277, 1443] 중에서 하나를 선택하는데, D2에서 14에 해당하는 리스트가 3488과 277을 포함하고 있으므로, 3488과 277과의 평균 거리가 가장 먼 샘플을 선택한다. 만약, 1443이 선택되었다고 하면 D2의 14에 해당하는 리스트에 1443을 삽입한다.
- [0058] 일 실시예에 따르면, L1에서 다섯 번째 요소인 14가 선택되고, D1에서 14에 해당하는 [3488, 5322, 277, 1443] 중에서 하나를 선택하는데, D2에서 14에 해당하는 리스트가 3488, 277, 1443을 포함하고 있어 남은 원소는 5322 하나뿐이므로 이를 선택한다.
- [0059] 일 실시예에 따르면, L1에서 여섯 번째 요소인 14가 선택되고, D1에서 14에 해당하는 [3488, 5322, 277, 1443] 중에서 하나를 선택하는데, D2에서 14에 해당하는 리스트가 3488, 277, 1443, 5322를 포함하고 있어 선택할 원소가 존재하지 않는다. 따라서, D2에서 14에 해당하는 리스트를 초기화하고 무작위로 선택한다. 만약, 277이 선

택되었다고 하면 D2의 14에 해당하는 리스트에 277을 삽입한다.

- [0060] 최종적으로 다수의 중복을 포함한 군집 번호 리스트(L1) 중 하나인 [14, 32, 14, 14, 14, 14]에 기반하여 샘플 번호 리스트(L2) [3488, 4222, 277, 1443, 5322, 277]이 생성되었다.
- [0061] 악성코드 탐지 모델(111)의 모델 관리부(250)는 샘플 추출부(240)가 신규 추출 및 추가 추출한 샘플들의 특징 벡터에 기반하여 악성코드 탐지 모델(111)을 앙상블 학습할 수 있다. 예를 들어, 모델 관리부(250)는 앙상블 학습한 이후, 악성코드 탐지 모델(111) 중 특정 샘플에 과적합 된 모델을 제거할 수 있다. 예를 들어, 모델 관리부(250)는 악성코드 탐지 모델(111)을 대상으로 가장 높은 중요도를 가지는 패턴을 추출하여 해당 패턴이 학습 셋에 포함되지 않은 샘플들로 구성된 검증셋(201)에 조회했을 때 존재하지 않으면 제거할 수 있다(S430).
- [0062] 모델 관리부(250)는 샘플 추출부(240)에서 출력된 각 샘플 번호 리스트(L2)에 포함된 샘플들의 특징 벡터(V)들에 기반하여 모델들을 학습하고, 각 모델을 대상으로 특정 샘플에 과적합 되었는지 검사하여 과적합 되었다고 판단된 모델은 제거할 수 있다. 결과적으로 일반화된 모델만을 학습한 앙상블 학습을 수행할 수 있다. 구체적으로는 샘플 추출부(240)에서 출력된 샘플 번호 리스트(L2)에 포함된 샘플 번호를 특징 정의부(210)에서 출력된 특징 벡터(V)에 조회하고, 조회된 특징 벡터(V)들에 기반하여 모델을 학습하는 것을 각 샘플 번호 리스트(L2)에 대하여 반복하여 다수의 모델을 학습하고, 학습된 모델들을 대상으로 가장 높은 중요도를 가지는 패턴을 추출하여 해당 패턴을 학습셋에 포함되지 않은 샘플들로 구성된 검증셋(201)에 조회했을 때 존재하지 않으면 제거한다.
- [0063] 한편, 전술한 실시예에 따른 악성코드 탐지 장치가 수행하는 학습 방법에 포함된 각각의 단계는, 이러한 단계를 프로세서가 수행하도록 하기 위한 명령어를 포함하여 기록매체에 기록된 컴퓨터 프로그램으로 구현될 수 있다.
- [0064] 또한, 전술한 실시예에 따른 대화 요악성코드 탐지 장치가 수행하는 학습 방법에 포함된 각각의 단계는, 이러한 단계를 프로세서가 수행하도록 하기 위한 명령어를 포함하는 컴퓨터 프로그램이 기록된 컴퓨터 판독 가능한 기록매체에서 구현될 수 있다.
- [0065] 본 발명에 첨부된 각 흐름도의 각 단계의 조합들은 컴퓨터 프로그램 인스트럭션들에 의해 수행될 수도 있다. 이들 컴퓨터 프로그램 인스트럭션들은 범용 컴퓨터, 특수용 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비의 프로세서에 탑재될 수 있으므로, 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비의 프로세서를 통해 수행되는 그 인스트럭션들이 흐름도의 각 단계에서 설명된 기능들을 수행하는 수단을 생성하게 된다. 이들 컴퓨터 프로그램 인스트럭션들은 특정 방식으로 기능을 구현하기 위해 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비를 지향할 수 있는 컴퓨터 이용 가능 또는 컴퓨터 판독 가능 기록매체에 저장되는 것도 가능하므로, 그 컴퓨터 이용가능 또는 컴퓨터 판독 가능 기록매체에 저장된 인스트럭션들은 흐름도의 각 단계에서 설명된 기능을 수행하는 인스트럭션 수단을 내포하는 제조 품목을 생산하는 것도 가능하다. 컴퓨터 프로그램 인스트럭션들은 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비 상에 탑재되는 것도 가능하므로, 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비 상에서 일련의 동작 단계들이 수행되어 컴퓨터로 실행되는 프로세스를 생성해서 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비를 수행하는 인스트럭션들은 흐름도의 각 단계에서 설명된 기능들을 실행하기 위한 단계들을 제공하는 것도 가능하다.
- [0066] 또한, 각 단계는 특정된 논리적 기능(들)을 실행하기 위한 하나 이상의 실행 가능한 인스트럭션들을 포함하는 모듈, 세그먼트 또는 코드의 일부를 나타낼 수 있다. 또, 몇 가지 대체 실시예들에서는 단계들에서 언급된 기능들이 순서를 벗어나서 발생하는 것도 가능함을 주목해야 한다. 예컨대, 잇달아 도시되어 있는 두 개의 단계들은 사실 실질적으로 동시에 수행되는 것도 가능하고 또는 그 단계들이 때때로 해당하는 기능에 따라 역순으로 수행되는 것도 가능하다.
- [0067] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 품질에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 발명에 개시된 실시예들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 균등한 범위 내에 있는 모든 기술사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

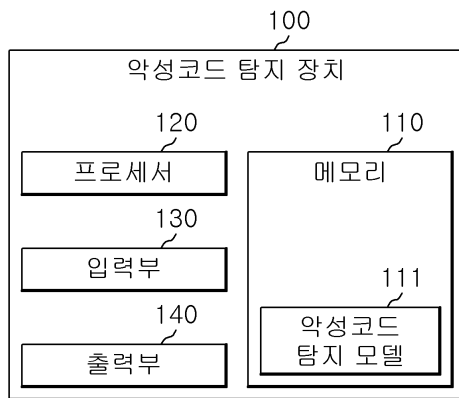
부호의 설명

- [0068] 100: 악성코드 탐지 장치

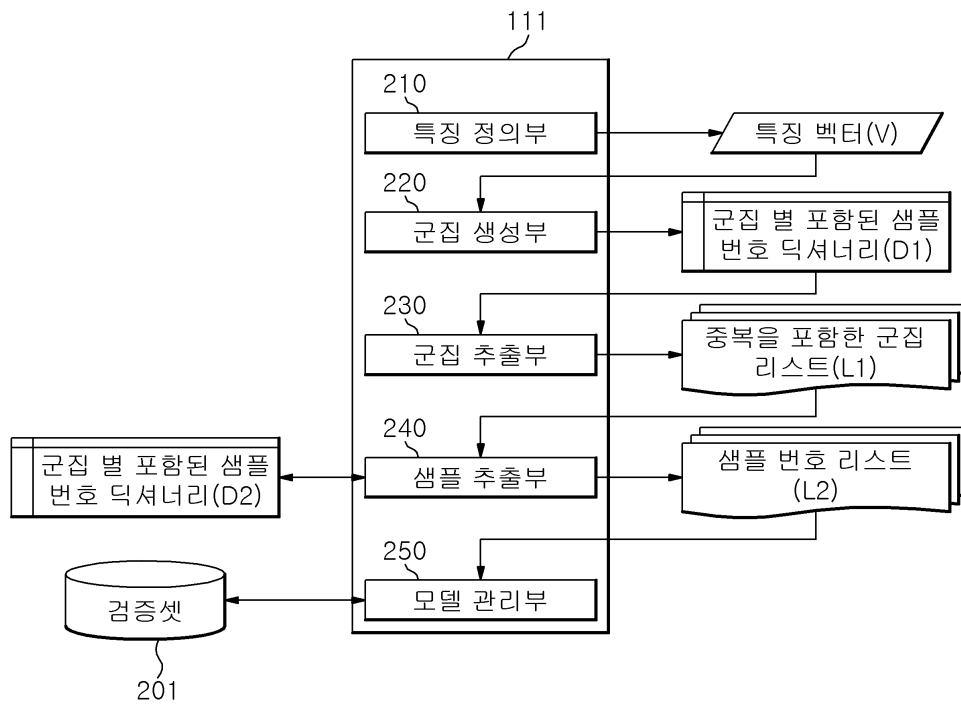
- 110: 메모리
- 111: 악성코드 탐지 모델
- 120: 프로세서
- 130: 입력부
- 140: 출력부

도면

도면1



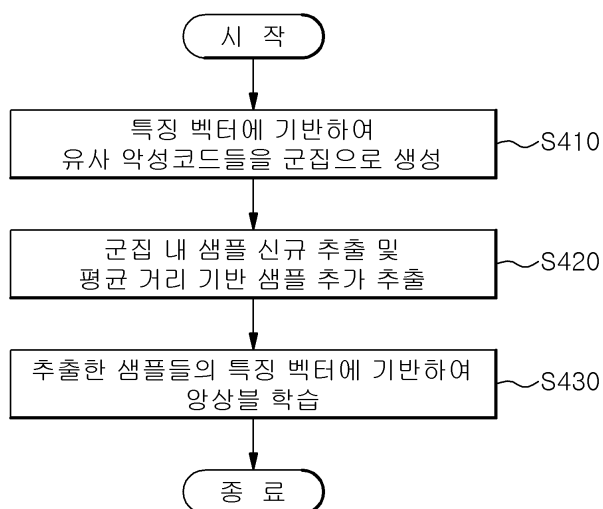
도면2



도면3



도면4



도면5

