(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0096226
(43) 공개일자 2022년07월07일

(51) 국제특허분류(Int. Cl.)
G06F 21/55 (2013.01) *G06F 21/56* (2013.01)
(52) CPC특허분류
G06F 21/554 (2013.01)
G06F 21/56 (2013.01)
(21) 출원번호 10-2020-0188483
(22) 출원일자 2020년12월30일
심사청구일자 2020년12월30일

(71) 출원인
주식회사 안랩
경기도 성남시 분당구 판교역로 220 (삼평동)
(72) 발명자
김원혁
경기도 성남시 수정구 제일로197번길 14-11, 2층 (태평동)
황용석
경기도 성남시 분당구 수내로 74, 116동 204호(수내동, 양지마을)
전제민
경기도 의왕시 갈미로 8, 204동 505호(내손동, 대원칸타빌2단지)
(74) 대리인
제일특허법인(유)

전체 청구항 수 : 총 11 항

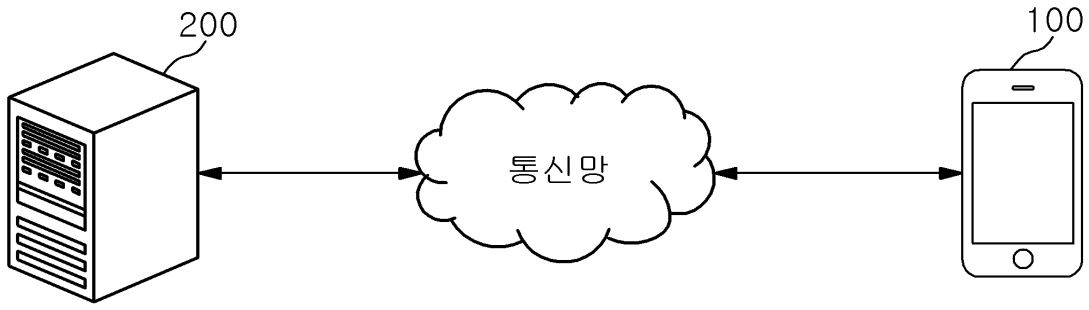
(54) 발명의 명칭 **화이트 리스트 생성 방법 및 이를 수행하는 사용자 단말, 컴퓨터 판독 가능한 기록 매체 및 컴퓨터 프로그램**

(57) 요 약

실시예는 사용자 단말에서 수행되는 화이트리스트 생성 방법에 있어서, 데이터를 수집하는 단계와, 상기 수집된 데이터의 악성 행위를 분석하고, 상기 분석된 데이터를 n차원 공간 상에 표시하는 잠재 공간 벡터를 출력하는 단계와, 상기 수집된 데이터 중 오탐 데이터를 탐지하는 단계와, 상기 악성 행위가 탐지된 데이터 및 상기 오탐 데이터를 기초로 화이트 리스트 모듈을 생성하는 단계를 포함할 수 있다.

실시예는 오탐 정보를 제공함으로써, 인공 신경망의 추가적인 학습없이 오탐을 개선할 수 있는 효과가 있다.

대 표 도 - 도1



명 세 서

청구범위

청구항 1

사용자 단말에서 수행되는 화이트리스트 생성 방법에 있어서,

데이터를 수집하는 단계;

상기 수집된 데이터의 악성 행위를 분석하고, 상기 분석된 데이터를 n차원 공간 상에 표시하는 잠재 공간 벡터를 출력하는 단계;

상기 악성 행위가 분석된 데이터 중 오탐 데이터를 탐지하는 단계; 및

상기 악성 행위가 탐지된 데이터 및 상기 오탐 데이터를 기초로 화이트 리스트 모듈을 생성하는 단계;

를 포함하는 화이트 리스트 생성 방법.

청구항 2

제1항에 있어서,

상기 화이트 리스트 모듈을 생성하는 단계는,

상기 잠재 공간 벡터 상의 상기 화이트 리스트 범위에 미리 설정된 임계값을 곱하여, 상기 화이트 리스트와 일부 중첩되는 다른 화이트 리스트들을 클러스터링하여 상기 화이트 리스트 모듈을 생성하는 화이트 리스트 생성 방법.

청구항 3

제1항에 있어서,

상기 화이트 리스트 모듈을 생성하는 단계는,

상기 잠재 공간 벡터 상의 서로 인접하는 화이트 리스트들 사이의 거리가 기준값 미만이면 인접하는 화이트 리스트들을 클러스터링하여 화이트 리스트 모듈을 생성하는 화이트 리스트 생성 방법.

청구항 4

제3항에 있어서,

상기 화이트 리스트 모듈을 생성하는 단계는,

상기 화이트 리스트들 중 대표 화이트 리스트를 설정하고, 상기 대표 화이트 리스트를 기준으로 다른 화이트 리스트들이 포함되는 상기 대표 화이트 리스트의 범위값을 설정하는 화이트 리스트 생성 방법.

청구항 5

제1항에 있어서,

상기 화이트 리스트 모듈을 생성하는 단계는,

상기 잠재 공간 벡터 상에 상기 오탐 데이터의 위치를 확인하고, 상기 오탐 데이터의 주변을 탐색하여 상기 오탐 데이터와 가장 가까운 악성 데이터와의 거리를 상기 화이트 리스트의 범위로 설정하여 상기 화이트 리스트 모듈을 생성하는 화이트 리스트 생성 방법.

청구항 6

제1항에 있어서,

상기 화이트 리스트 모듈을 생성하는 단계는,

상기 잠재 공간 벡터를 신규 인공 신경망의 공간 벡터로 이용하는 화이트 리스트 생성 방법.

청구항 7

제1항에 있어서,

상기 오탐 데이터를 탐지하는 단계에서,

상기 오탐 데이터는 검색 엔진, 일반 사용자 또는 분석가로부터 수집되는 화이트 리스트 생성 방법.

청구항 8

제1항에 있어서,

상기 데이터를 수집하는 단계에서,

상기 데이터는 안티 바이러스 프로그램에 의해 수집되는 화이트 리스트 생성 방법.

청구항 9

화이트 리스트를 생성하는 단말에 있어서,

화이트 생성 프로그램이 저장된 메모리; 및

상기 메모리에 저장된 프로그램을 실행하는 프로세서를 포함하고,

상기 프로세서는

데이터를 수집하고, 상기 수집된 데이터의 악성 행위를 분석하고, 상기 분석된 데이터를 n차원 공간 상에 표시하는 잠재 공간 벡터를 출력하고, 상기 악성 행위가 분석된 데이터 중 오탐 데이터를 탐지하고, 상기 악성 행위가 탐지된 데이터 및 상기 오탐 데이터를 기초로 화이트 리스트 모듈을 생성하는 화이트 리스트를 생성하는 단말.

청구항 10

컴퓨터 프로그램을 저장하고 있는 컴퓨터 판독 가능 기록매체로서,

상기 컴퓨터 프로그램은, 프로세서에 의해 실행되면,

제1항 내지 제8항 중 어느 한 항에 따른 방법을 포함하는 동작을 상기 프로세서가 수행하도록 하기 위한 명령어를 포함하는 컴퓨터 판독 가능한 기록매체.

청구항 11

컴퓨터 판독 가능한 기록매체에 저장되어 있는 컴퓨터 프로그램으로서,

상기 컴퓨터 프로그램은, 프로세서에 의해 실행되면,

제1항 내지 제8항 중 어느 한 항에 따른 방법을 포함하는 동작을 상기 프로세서가 수행하도록 하기 위한 명령어를 포함하는 컴퓨터 프로그램.

발명의 설명

기 술 분 야

[0001] 실시예는 화이트 리스트를 생성하기 위한 화이트 리스트 생성 장치 및 방법에 관한 것이다.

배 경 기 술

[0002] 일반적으로, 화이트 리스트에 등록하여 관리하는 방법은 사용자의 주관적인 판단 및 디지털 서명과 같은 단순한 방법으로 화이트 리스트에 등록한다. 즉 종래의 화이트리스 등록 방법은 기존에 사용하던 소프트웨어이거나 안티 바이러스 제품을 이용하여 검출되지 않은 것들을 화이트 리스트에 등록한다.

[0003] 종래의 화이트 리스트 등록 방법은 블랙 리스트를 기반의 악성 코드 탐지하는 엔진을 이용하고 있으나, 100% 정

확도를 가지는 엔진은 개발할 수 없기 때문에 오탐이 발생하게 된다. 이러한 오탐으로 인해 서비스 편의성을 해칠 수 있게 된다.

[0004] 예를 들어, 만약 악성 행위를 발견해 로그를 띄운다고 가정할 때, 윈도우 프로그램을 오탐하는 경우, 계속 로그를 띄워 문제를 야기할 수 있다. 이러한 경우, 프로그램을 삭제하여 문제를 해소할 수 있으나, 이는 또 다른 문제점을 발생시킬 수 있다.

[0005] 또한, 서비스 환경에 따라 오탐으로 등장하는 악성 패턴의 종류도 다양하게 등장하는데 모든 환경의 오탐에 대응하는 화이트 리스트를 만드는 것은 현실적으로 불가능하다.

[0006] 또한, 데이터는 프로그램의 행위 데이터를 사용하여 악성 행위를 탐지하는 데 행위 패턴을 분석하여 정상, 악성으로 판단하는 것은 분석가들이 시간을 많이 사용하여 분석을 해야하고, 패턴에 대한 아주 세밀한 조건들을 설정해야 하기 때문에 분석을 통한 예외 처리를 진행하는 것도 불가능하다.

발명의 내용

해결하려는 과제

[0007] 상술한 문제점을 해결하기 위해, 실시예는 화이트 리스트 생성시 별도의 오탐으로 탐지된 오탐 데이터를 이용하여 정확성을 향상시키기 위한 화이트 리스트 생성 장치 및 방법을 제공하는 것을 그 목적으로 한다.

과제의 해결 수단

[0008] 실시예는 사용자 단말에서 수행되는 화이트리스트 생성 방법에 있어서, 데이터를 수집하는 단계와, 상기 수집된 데이터의 악성 행위를 분석하고, 상기 분석된 데이터를 n차원 공간 상에 표시하는 잠재 공간 벡터를 출력하는 단계와, 상기 악성 행위가 분석된 데이터 중 오탐 데이터를 탐지하는 단계와, 상기 악성 행위가 탐지된 데이터 및 상기 오탐 데이터를 기초로 화이트 리스트 모듈을 생성하는 단계를 포함할 수 있다.

[0009] 상기 화이트 리스트 모듈을 생성하는 단계는, 상기 잠재 공간 벡터 상의 상기 화이트 리스트 범위에 미리 설정된 임계값을 곱하여, 상기 화이트 리스트와 일부 중첩되는 다른 화이트 리스트들을 클러스터링하여 상기 화이트 리스트 모듈을 생성할 수 있다.

[0010] 상기 화이트 리스트 모듈을 생성하는 단계는, 상기 잠재 공간 벡터 상의 서로 인접하는 화이트 리스트들 사이의 거리가 기준값 미만이면 인접하는 화이트 리스트들을 클러스터링하여 화이트 리스트 모듈을 생성할 수 있다.

[0011] 상기 화이트 리스트 모듈을 생성하는 단계는, 상기 화이트 리스트들 중 대표 화이트 리스트를 설정하고, 상기 대표 화이트 리스트를 기준으로 다른 화이트 리스트들이 포함되는 상기 대표 화이트 리스트의 범위값을 설정할 수 있다.

[0012] 상기 화이트 리스트 모듈을 생성하는 단계는, 상기 잠재 공간 벡터 상에 상기 오탐 데이터의 위치를 확인하고, 상기 오탐 데이터의 주변을 탐색하여 상기 오탐 데이터와 가장 가까운 악성 데이터와의 거리를 상기 화이트 리스트의 범위로 설정하여 상기 화이트 리스트 모듈을 생성할 수 있다.

[0013] 상기 화이트 리스트 모듈을 생성하는 단계는, 상기 잠재 공간 벡터를 신규 인공 신경망의 공간 벡터로 이용할 수 있다.

[0014] 상기 오탐 데이터를 탐지하는 단계에서, 상기 오탐 데이터는 검색 엔진, 일반 사용자 또는 분석가로부터 수집될 수 있다.

[0015] 상기 데이터를 수집하는 단계에서, 상기 데이터는 안티 바이러스 프로그램에 의해 수집될 수 있다.

[0016] 또한, 실시예는 화이트 리스트를 생성하는 단말에 있어서, 화이트 생성 프로그램이 저장된 메모리와, 상기 메모리에 저장된 프로그램을 실행하는 프로세서를 포함하고, 상기 프로세서는 데이터를 수집하고, 상기 수집된 데이터의 악성 행위를 분석하고, 상기 분석된 데이터를 n차원 공간 상에 표시하는 잠재 공간 벡터를 출력하고, 상기 악성 행위가 분석된 데이터 중 오탐 데이터를 탐지하고, 상기 악성 행위가 탐지된 데이터 및 상기 오탐 데이터를 기초로 화이트 리스트 모듈을 생성할 수 있다.

[0017] 또한, 실시예는 컴퓨터 프로그램을 저장하고 있는 컴퓨터 판독 가능 기록매체로서, 상기 컴퓨터 프로그램은, 프로세서에 의해 실행되면, 데이터를 수집하고, 상기 수집된 데이터의 악성 행위를 분석하고, 상기 분석된 데이터를 n차원 공간 상에 표시하는 잠재 공간 벡터를 출력하고, 상기 악성 행위가 분석된 데이터 중 오탐 데이터를

탐지하고, 상기 악성 행위가 탐지된 데이터 및 상기 오탐 데이터를 기초로 화이트 리스트 모듈을 생성하는 동작을 상기 프로세서가 수행하도록 하기 위한 명령어를 포함할 수 있다.

[0018] 또한, 실시예는 컴퓨터 판독 가능한 기록매체에 저장되어 있는 컴퓨터 프로그램으로서, 상기 컴퓨터 프로그램은, 프로세서에 의해 실행되면, 데이터를 수집하고, 상기 수집된 데이터의 악성 행위를 분석하고, 상기 분석된 데이터를 n차원 공간 상에 표시하는 잠재 공간 벡터를 출력하고, 상기 악성 행위가 분석된 데이터 중 오탐 데이터를 탐지하고, 상기 악성 행위가 탐지된 데이터 및 상기 오탐 데이터를 기초로 화이트 리스트 모듈을 생성하는 동작을 상기 프로세서가 수행하도록 하기 위한 명령어를 포함할 수 있다.

발명의 효과

[0019] 실시예는 오탐 정보를 제공함으로써, 인공 신경망의 추가적인 학습없이 오탐을 개선할 수 있는 효과가 있다.

[0020] 또한, 실시예는 사용자가 탐지한 오탐 정보를 제공함으로써, 일반 사용자 환경에 맞도록 오탐을 개선할 수 있는 효과가 있다.

[0021] 또한, 실시예는 분석가들의 업무를 지원하여 더 많은 양의 데이터들을 분석하여 엔진의 성능을 향상시킬 수 있는 효과가 있다.

도면의 간단한 설명

[0022] 도 1은 실시예에 따른 화이트 리스트 생성 시스템의 구성을 나타낸 도면이다.

도 2는 실시예에 따른 화이트 리스트 생성 장치가 구비된 사용자 단말의 구성을 나타낸 도면이다.

도 3은 실시예에 따른 화이트 리스트 생성 방법을 나타낸 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0023] 이하, 도면을 참조하여 실시예를 상세히 설명하기로 한다.

[0025] 도 1은 실시예에 따른 화이트 리스트 생성 시스템의 구성을 나타낸 도면이고, 도 2는 실시예에 따른 화이트 리스트 생성 장치가 구비된 사용자 단말의 구성을 나타낸 도면이다.

[0026] 도 1을 참조하면, 실시예에 따른 화이트 리스트 생성 시스템은 사용자 단말(100)과, 서버(200)를 포함하고, 통신망을 통해 상호 연결될 수 있다.

[0027] 사용자 단말(100)은 사용자가 사용하는 단말로서 인공 신경망 및 오탐 데이터를 이용하여 화이트 리스트를 생성하는 단말을 의미하며, 서버(200)는 화이트 리스트 공유를 위한 장치일 수 있다.

[0028] 사용자 단말(100)은 저장중인 소프트웨어 파일을 인식하고, 소프트웨어 파일로부터 속성 정보를 추출하게 된다. 사용자 단말(100)은 속성 정보에 기초하여 소프트웨어 파일에 대해 인공 신경망 및 오탐 데이터를 이용하여 화이트 리스트를 생성하게 된다.

[0029] 사용자 단말(100)은 생성된 화이트 리스트를 서버(200)로 공유하거나 다른 사용자 단말에 전송하여 다른 사용자 단말에 저장된 화이트 리스트를 업데이트할 수 있다.

[0030] 속성 정보란 정적 데이터, 동적 데이터, 휴리스틱 데이터 및 사회공학적 데이터 중 적어도 하나 이상을 포함하는 정보로서 이를 상세하게 서술하면 아래와 같다.

[0031] 정적 데이터는 검증 대상 소프트웨어의 디지털 서명 보유 유무, 파일 설명(description) 보유 유무, 파일 버전의 최신 여부, 제품 버전의 최신 여부, 제품 이름 보유유무, 저작권 보유 유무 중 적어도 하나를 포함하는 정보일 수 있다.

[0032] 동적 데이터는 검증 대상 소프트웨어가 실행 시 산출되는 프로세스 정보, 메모리 상태, 설치 경로, 검증 대상 소프트웨어의 실행 및 사용 이력, 등록현황 및 검증 대상 소프트웨어의 실행 상태 정보 중 적어도 하나를 포함하는 정보일 수 있다.

[0033] 휴리스틱 데이터는 검증 대상 소프트웨어의 메모리, 용량, CPU점유율, 네트워크, 파일 입출력, 스레드 사용량 및 임계치 증가율을 포함하는 비정상 행위에 대한 데이터 중 적어도 하나를 포함하는 정보일 수 있다

[0034] 사회공학적 데이터는 검증 대상 소프트웨어의 입력장치에 의한 적어도 한번 이상의 실행 여부, 파일 이름, 제품 이름 및 저작권 정보의 해석 가능 여부, 실행 시간의 기준 기간 경과 여부, 사용률이 기준 값 이상 인지 여부, 및 기준 퍼센테이지 이상의 사용자 인증 여부 중 적어도 하나를 포함하는 정보일 수 있다.

[0035] 사용자 단말(100)은 스마트폰, 컴퓨터, 노트북 등의 어플리케이션을 이용할 수 있는 통신 단말기일 수 있으나, 그 종류는 한정되지 않는다.

[0036] 또한, 사용자 단말(100)은 핸드헬드 컴퓨팅 디바이스(예를 들면, PDA, 이메일 클라이언트 등), 핸드폰의 임의의 형태, 또는 다른 종류의 컴퓨팅 또는 커뮤니케이션 플랫폼의 임의의 형태를 포함할 수 있으나, 이에 한정되는 것은 아니다.

[0037] 서버(200)는 화이트 리스트를 공유하기 위한 서버로서, 다른 사용자 단말을 통해 설정된 화이트 리스트를 서버(200)를 통해 공유할 수 있게 된다.

[0038] 통신망은 사용자 단말(100)과 서버(200)들을 연결하는 역할을 수행한다. 즉, 통신망은 사용자 단말(100)들이 서버(200)에 접속한 후 데이터를 송수신할 수 있도록 접속 경로를 제공하는 통신망을 의미한다. 통신망은 예컨대 LANs(Local Area Networks), WANs(Wide Area Networks), MANs(Metropolitan Area Networks), ISDNs(Integrated Service Digital Networks) 등의 유선 네트워크나, 무선 LANs, CDMA, 블루투스, 위성 통신 등의 무선 네트워크를 망라할 수 있으나, 이에 한정되는 것은 아니다.

[0039] 도 2를 참조하면, 사용자 단말(100)은 통신 모듈(110), 메모리(120) 및 프로세서(130)를 포함할 수 있다.

[0040] 통신 모듈(110)은 통신망과 연동하여 사용자 단말(100)과 서버(200) 간의 송수신 신호를 패킷 데이터 형태로 제공하는 데 필요한 통신 인터페이스를 제공한다. 나아가, 통신 모듈(110)은 서버(200)로부터 데이터 요청을 수신하고, 이에 대한 응답으로서 데이터를 송신하는 역할을 수행할 수 있다.

[0041] 통신 모듈(110)은 다른 네트워크 장치와 유무선 연결을 통해 제어 신호 또는 데이터 신호와 같은 신호를 송수신하기 위해 필요한 하드웨어 및 소프트웨어를 포함하는 장치일 수 있다.

[0042] 메모리(120)는 화이트 리스트 생성을 수행하기 위한 프로그램이 기록된다. 또한, 프로세서(130)가 처리하는 데이터를 일시적 또는 영구적으로 저장하는 기능을 수행한다. 여기서, 메모리(120)는 자기 저장 매체(magnetic storage media) 또는 플래시 저장 매체(flash storage media)를 포함할 수 있으나, 이에 한정되는 것은 아니다.

[0043] 메모리(120)에는 화이트 리스트를 포함한 화이트 리스트 생성에 필요한 정보가 저장될 수 있다. 이는 예컨대, 소프트웨어 파일과 상기 소프트웨어 파일에서 추출된 속성 정보 및 학습이 수행된 인공 신경망 및 오탐 데이터에 대한 정보가 추가로 저장될 수 있다.

[0044] 프로세서(130)는 일종의 중앙처리장치로서 인공 신경망 및 오탐 데이터를 이용하여 화이트 리스트를 생성하는 전체 과정을 제어할 수 있다. 프로세서가 수행하는 각 단계는 도 3을 참조하여 상세히 설명하기로 한다.

[0045] 여기서, 프로세서(130)는 프로세서(processor)와 같이 데이터를 처리할 수 있는 모든 종류의 장치를 포함할 수 있다. 여기서, '프로세서(processor)'는, 예를 들어 프로그램 내에 포함된 코드 또는 명령으로 표현된 기능을 수행하기 위해 물리적으로 구조화된 회로를 갖는, 하드웨어에 내장된 데이터 처리 장치를 의미할 수 있다. 이와 같이 하드웨어에 내장된 데이터 처리 장치의 일 예로써, 마이크로프로세서(microprocessor), 중앙처리장치(central processing unit: CPU), 프로세서 코어(processor core), 멀티프로세서(multiprocessor), ASIC(application-specific integrated circuit), FPGA(field programmable gate array) 등의 처리 장치를 망라할 수 있으나, 이에 한정되는 것은 아니다.

[0046] 도 3는 실시예에 따른 화이트 리스트 생성 방법을 나타낸 순서도이다.

[0047] 도 3을 참조하면, 실시예에 따른 화이트 리스트 생성 방법은 데이터를 수집하는 단계(S100)를 수행할 수 있다.

[0048] 단계(S100)에서 데이터는 사용자 단말에서 실행되는 소프트웨어 파일을 인식하고, 그에 대한 속성 정보일 수 있다. 속성 정보는 정적 데이터, 동적 데이터, 휴리스틱 데이터 및 사회공학적 데이터 중 어느 하나 이상을 포함할 수 있으나, 이에 한정되지 않는다.

[0049] 단계(S100)에서 데이터는 안티 바이러스 프로그램 예컨대, V3를 통해 수집되는 데이터일 수 있다.

[0050] 단계(S100)에서 수집된 데이터는 메모리에 저장될 수 있다.

[0051] 실시예에 따른 화이트 리스트 생성 방법은 수집된 데이터를 악성 행위를 분석하고, 이를 n차원 공간 상에 표시하는 잠재 공간 벡터를 출력하는 단계(S200)를 수행할 수 있다. 여기서, n 차원 공간은 1차원, 2차원 또는 3차원의 공간일 수 있으나, 이에 한정되지 않는다.

[0052] 단계(S200)에서 악성 행위 분석 및 잠재 공간 벡터는 메모리에 저장된 인공 신경망을 이용하여 출력할 수 있다. 인공 신경망은 딥러닝 알고리즘, 머신러닝 알고리즘, 지도학습 알고리즘, 비지도 학습 알고리즘 중 어느 하나를 이용할 수 있으며 그 종류는 한정되지 않는다.

[0053] 인공 신경망은 메모리에 저장된 데이터를 입력받아 악성 행위를 분석할 수 있다. 인공 신경망은 메모리에 저장된 데이터를 입력받아 분석된 데이터를 2차원 공간 상에 표시하는 잠재 공간 벡터를 출력할 수 있다.

[0054] 실시예에 따른 화이트 리스트 생성 방법은 오탐 데이터를 탐지하는 단계(S300)를 수행할 수 있다.

[0055] 단계(S300)에서 악성 행위가 분석된 데이터를 이용하여 오탐 데이터를 탐지할 수 있다. 오탐 데이터는 검색 엔진, 사용자 또는 분석가로부터 분석되어 수집할 수 있다.

[0056] 단계(S300)에서 오탐이라고 탐지된 오탐 데이터는 메모리에 저장될 수 있다.

[0057] 실시예에 따른 화이트 리스트 생성 방법은 악성 행위가 탐지된 데이터 및 오탐 데이터를 기초로 화이트 리스트 모듈을 생성하는 단계(S400)를 수행할 수 있다.

[0058] 단계(S400)에서 클러스터링 기법을 이용하여 화이트 리스트 모듈을 생성 및 최적화할 수 있다. 모든 오탐 데이터를 화이트 리스트로 생성할 경우, 비교해야 할 데이터의 양이 많기 때문에 연산 시간이 느려지고 메모리 사용량이 많아질 수 있다. 이에 클러스터링 기법을 이용하여 최적화를 수행할 수 있다.

[0059] 일 예로, 단계(S400)에서 먼저, 인공 신경망에서 출력된 잠재 공간 벡터 상의 화이트 리스트에 미리 설정된 임계값을 곱하여 화이트 리스트 범위를 설정할 수 있다.

[0060] 미리 설정된 임계값을 곱하여 화이트 리스트 범위를 설정하는 이유는 데이터가 이미 화이트 리스트 범위에 포함되어 있는 경우, 화이트 처리되어 새로 생성되지 않는다. 따라서, 잠재 공간 벡터 상의 다른 화이트 리스트가 화이트 리스트 범위에 일정 부분 이상 겹치는 경우 하나의 클러스터로 묶기 위함이다.

[0061] 다른 예로, 단계(S400)에서 잠재 공간 벡터 상의 화이트 리스트들의 거리 값을 이용하여 서로 인접하는 화이트 리스트들을 클러스터링하여 화이트 리스트 모듈을 생성할 수 있다.

[0062] 예를 들어, 잠재 공간 벡터 상의 화이트 리스트 범위가 4인 벡터 A와 잠재 공간 벡터 상의 화이트 리스트 범위가 2인 벡터 B가 있고, 벡터 A와 벡터 B의 사이가 3일 때, 거리가 임계값 미만을 조건을 충족할 경우, 벡터 B는 벡터 A의 범위에 포함될 수 있도록 클러스터링할 수 있다.

[0063] 클러스터링이 끝나면 화이트 리스트들 중 가장 영향력이 있는 화이트 리스트 또는 중간 벡터를 대표 화이트 리스트로 설정하고, 대표 화이트 리스트를 기준으로 다른 화이트 리스트들이 포함될 수 있도록 대표 화이트 리스트의 범위값을 설정할 수 있다.

[0064] 또 다른 예로, 단계(S400)에서 먼저, 잠재 공간 벡터 상에 오탐 데이터를 위치를 확인할 수 있다. 이어서, 가장 가까운 데이터 N개(사용자 정의 값)을 지정하여 주변에 정상 데이터가 많은 지 악성 데이터가 많은 지 탐색할 수 있다. 이어서, 가장 가까운 악성 데이터와의 거리를 화이트 리스트의 범위로 설정하여 화이트 리스트 모듈을 생성할 수 있다.

[0065] 만약 악성 데이터가 많은 경우, 라벨(Label) 노이즈에 의해 잘못된 라벨 일 수 있기 때문에 이러한 데이터가 화이트 리스트로 처리되지 않게 하기 위해 주변을 탐색하고 가장 가까운 데이터들이 N% 이상 정상 데이터이면 화이트 리스트를 생성할 준비를 할 수 있다.

[0066] 화이트 리스트 범위가 너무 큰 경우, 실제 악성 데이터를 포함할 위험성이 있기 때문에 최소 악성 데이터와의 거리를 이용하여 설정하게 되면 악성 데이터를 포함하지 않는 화이트 리스트 모듈을 생성할 수 있게 된다.

[0067] 또 다른 예로, 단계(S400)에서 오탐으로 판단된 잠재 공간 벡터와 사용자 정의 화이트 리스트 범위를 사용하여 화이트 리스트 모듈을 생성할 수 있다. 사용자 정의를 사용하는 이유는 최초 화이트 리스트를 생성할 때는 가장 가까운 악성 데이터를 찾아 비율을 조회하지만 현재는 데이터가 하나밖에 없고, 화이트 리스트 밖에 정보가 없기 때문에 불가능하다, 따라서, 작은 값을 이용하여 화이트 리스트 생성에 사용할 수 있다.

[0068] 상기와 같이 생성된 화이트 리스트를 적용하여 클러스터링을 수행하고, 수행결과 생성된 화이트 리스트를 업데이트할 수 있다.

[0069] 또 다른 예로, 단계(S400)에서 인공 신공망으로부터 출력되는 잠재 벡터 공간 벡터를 메모리에 저장할 수 있다. 단계(S400)에서 새로운 신규 인공 신경망의 잠재 벡터 공간을 구하고, 2개의 잠재 벡터 공간 변환 레이어를 연산하여 신규 인공 신경망을 기존 잠재 벡터 공간으로 변형하는 레이어를 구할 수 있다.

[0070] 구해진 잠재 벡터 공간 변환 레이어를 신규 인공 신경망에 적용하고 기존 화이트 리스트를 재사용할 수 있다.

[0071] 상기와 같이, 화이트 리스트 모듈이 생성되면, 생성된 화이트 리스트 모듈을 적용하는 단계를 더 수행할 수도 있다. 예컨대 인공 신경망에서 악성으로 판단된 데이터일 경우, 잠재 공간 벡터를 화이트 리스트 적용 모듈에 받을 수 있다. 수집된 잠재 공간 벡터를 화이트 리스트에 조회하여 수집된 잠재 공간 벡터가 포함되어 있는지 확인할 수 있다.

[0072] 화이트 리스트에 포함되어 있으면 악성 결과를 정상으로 수정하여 사용자 단말의 UI 모듈로 전달할 수 있다. UI 모듈은 사용자 또는 서버와 상호 작용을 하는 모듈일 수 있다.

[0073] 실시예는 오탐 정보를 제공함으로써, 인공 신경망의 추가적인 학습없이 오탐을 개선할 수 있는 효과가 있다.

[0074] 또한, 실시예는 사용자가 탐지한 오탐 정보를 제공함으로써, 일반 사용자 환경에 맞도록 오탐을 개선할 수 있는 효과가 있다.

[0075] 또한, 실시예는 분석가들의 업무를 지원하여 더 많은 양의 데이터들을 분석하여 엔진의 성능을 향상시킬 수 있는 효과가 있다.

[0077] 본 문서의 다양한 실시예들은 기기(machine)(예: 컴퓨터)로 읽을 수 있는 저장 매체(machine-readable storage media)(예: 메모리(내장 메모리 또는 외장 메모리))에 저장된 명령어를 포함하는 소프트웨어(예: 프로그램)로 구현될 수 있다. 기기는, 저장 매체로부터 저장된 명령어를 호출하고, 호출된 명령어에 따라 동작이 가능한 장치로서, 개시된 실시예들에 따른 전자 장치를 포함할 수 있다. 상기 명령이 제어부에 의해 실행될 경우, 제어부가 직접, 또는 상기 제어부의 제어하에 다른 구성요소들을 이용하여 상기 명령에 해당하는 기능을 수행할 수 있다. 명령은 컴파일러 또는 인터프리터에 의해 생성 또는 실행되는 코드를 포함할 수 있다. 기기로 읽을 수 있는 저장매체는, 비일시적(non-transitory) 저장매체의 형태로 제공될 수 있다. 여기서, 비일시적은 저장매체가 신호(signal)를 포함하지 않으며 실재(tangible)한다는 것을 의미할 뿐 데이터가 저장매체에 반영구적 또는 임시적으로 저장됨을 구분하지 않는다.

[0078] 실시예에 따르면, 본 문서에 개시된 다양한 실시예들에 따른 방법은 컴퓨터 프로그램 제품(computer program product)에 포함되어 제공될 수 있다.

[0079] 일 실시예에 따르면, 컴퓨터 프로그램을 저장하고 있는 컴퓨터 판독 가능 기록매체로서, 데이터를 수집하고, 상기 수집된 데이터의 악성 행위를 분석하고, 상기 분석된 데이터를 n차원 공간 상에 표시하는 잠재 공간 벡터를 출력하고, 상기 수집된 데이터 중 오탐 데이터를 탐지하고, 상기 악성 행위가 탐지된 데이터 및 상기 오탐 데이터를 기초로 화이트 리스트 모듈을 생성하는 동작을 포함하는 방법을 프로세서가 수행하도록 하기 위한 명령어를 포함할 수 있다.

[0080] 일 실시예에 따르면, 컴퓨터 판독 가능한 기록매체에 저장되어 있는 컴퓨터 프로그램으로서, 데이터를 수집하고, 상기 수집된 데이터의 악성 행위를 분석하고, 상기 분석된 데이터를 n차원 공간 상에 표시하는 잠재 공간 벡터를 출력하고, 상기 수집된 데이터 중 오탐 데이터를 탐지하고, 상기 악성 행위가 탐지된 데이터 및 상기 오탐 데이터를 기초로 화이트 리스트 모듈을 생성하는 동작을 포함하는 방법을 프로세서가 수행하도록 하기 위한 명령어를 포함할 수 있다.

[0082] 상기에서는 도면 및 실시예를 참조하여 설명하였지만, 해당 기술 분야의 숙련된 당업자는 하기의 특허청구범위에 기재된 실시예의 기술적 사상으로부터 벗어나지 않는 범위 내에서 실시예는 다양하게 수정 및 변경시킬 수 있음은 이해할 수 있을 것이다.

부호의 설명

[0083] 100: 사용자 단말

110: 통신 단말

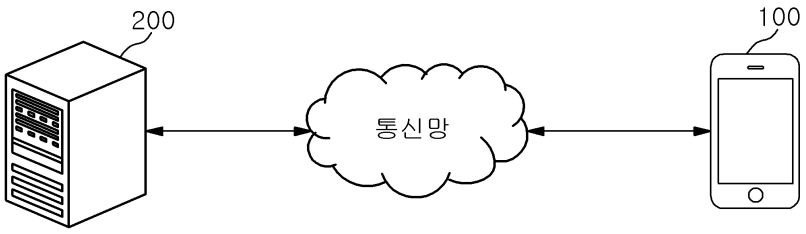
120: 메모리

130: 프로세서

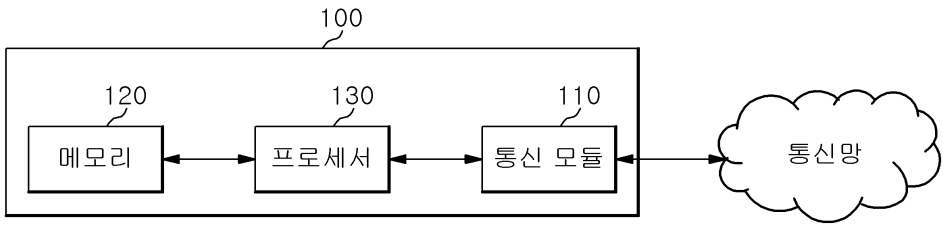
200: 서버

도면

도면1



도면2



도면3