



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2015년07월15일

(11) 등록번호 10-1537088

(24) 등록일자 2015년07월09일

(51) 국제특허분류(Int. Cl.)

G06F 21/10 (2013.01)

(21) 출원번호 10-2014-0116394

(22) 출원일자 2014년09월02일

심사청구일자 2014년09월02일

(56) 선행기술조사문헌

JP2012501028 A

KR1020120073018 A

JP2014071796 A

(73) 특허권자

인포섹(주)

서울특별시 강남구 영동대로 316, 성원빌딩 (대치동)

(72) 발명자

조래현

경기도 수원시 영통구 센트럴파크로 100, 6407동 1804호 (이의동, 광교 센트럴타운 오드카운티)

이동희

경기도 광명시 소하로 189, 207동 504호 (소하동, 신촌휴먼시아)

한인희

서울특별시 강남구 현릉로590길 100, 104동 903호 (세곡동, 리엔파크1단지)

(74) 대리인

강대훈, 나선균, 방영석

전체 청구항 수 : 총 19 항

심사관 : 문남두

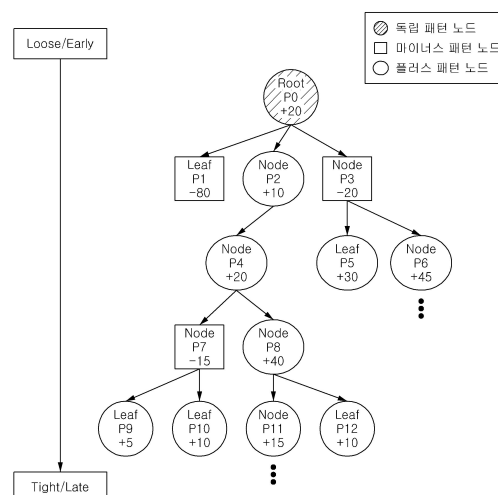
(54) 발명의 명칭 API 호출 흐름 기반의 악성코드 탐지 시스템 및 방법

(57) 요약

본 발명은 API 호출 흐름 기반의 악성코드 탐지 시스템 및 방법에 관한 것으로, 보다 상세하게는 프로그램 상의 보호된 영역인 샌드박스(Sandbox) 내에서 분석 대상 코드를 동작시켰을 때, API의 일련의 호출 흐름에 따른 행위 데이터를 분석하여 악성코드를 탐지하는 기술에 관한 것이다.

이러한 목적을 달성하기 위하여 본 발명의 일 실시예에 따른 API 호출 흐름 기반의 악성코드 탐지 시스템은 행위 데이터 수집부, 매칭부, 가중치 합산부 및 악성코드 판단부를 포함한다.

대표도 - 도6



명세서

청구범위

청구항 1

프로그램 상의 보호된 영역 내에서 분석 대상 코드를 동작시키는 단계;

상기 보호된 영역 내에서 상기 분석 대상 코드가 동작하는 동안 상기 분석 대상 코드가 호출하는 API(Application Programming Interface)의 일련의 호출 흐름에 따른 상기 분석 대상 코드의 행위 데이터를 수집하는 단계;

상기 수집된 행위 데이터를, 각 노드가 특정 행위에 대응하도록 형성된 트리 구조의 행위 패턴과 순차적으로 매칭하는 단계;

상기 수집된 행위 데이터와 순차적으로 매칭된 상기 행위 패턴 내의 각 노드에 부여된 가중치를 합산하는 단계; 및

상기 합산된 가중치가 기 정해진 합산 값 이상일 경우, 상기 분석 대상 코드를 악성코드로 판단하는 단계;

를 포함하는 API 호출 흐름 기반의 악성코드 탐지 방법.

청구항 2

제1항에 있어서,

상기 매칭하는 단계는

상기 트리 구조의 상기 행위 패턴 내의 현재 노드와 연결되는 다음 노드들 중, 상기 수집된 행위 데이터의 행위가 상기 현재 노드와 일치하는지의 여부에 따라 상기 다음 노드들과 순차적으로 매칭하는 것

을 특징으로 하는 API 호출 흐름 기반의 악성코드 탐지 방법.

청구항 3

제1항에 있어서,

상기 합산하는 단계는

악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치가 부여되고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치가 부여되는 상기 트리 구조의 상기 행위 패턴을 기반으로, 상기 행위 데이터와 매칭된 상기 행위 패턴 내의 상기 가중치를 합산하는 것

을 특징으로 하는 API 호출 흐름 기반의 악성코드 탐지 방법.

청구항 4

제1항에 있어서,

상기 수집하는 단계는

상기 분석 대상 코드가 동작하는 동안 프로세스 별로 생성되거나 또는 기록된 상기 행위 데이터를 시간 순으로 수집하는 것

을 특징으로 하는 API 호출 흐름 기반의 악성코드 탐지 방법.

청구항 5

프로그램 상의 보호된 영역 내에서 분석 대상 코드를 동작시키는 단계;

상기 보호된 영역 내에서 상기 분석 대상 코드가 동작하는 동안 상기 분석 대상 코드가 호출하는 API(Application Programming Interface)를 식별하고, 상기 식별된 API에 따라서 상기 분석 대상 코드의 행위

데이터를 수집하는 단계;

각 노드가 특정 행위에 대응하도록 형성된 트리 구조의 행위 패턴 내의 현재 노드와 연결되는 다음 노드들 중, 상기 수집된 상기 분석 대상 코드의 행위 데이터와 일치하는 행위에 대응하는 노드를 식별하여 매칭하는 단계;

상기 매칭된 노드에 부여된 가중치를 누적하여 합산하는 단계; 및

상기 합산된 가중치가 기 정해진 합산 값 이상일 경우, 상기 분석 대상 코드를 악성코드로 판단하는 단계;

를 포함하는 API 호출 흐름 기반의 악성코드 탐지 방법.

청구항 6

제5항에 있어서,

상기 합산하는 단계는

악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치가 부여되고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치가 부여된 상기 매칭된 노드의 상기 가중치를 누적하여 합산하는 것

을 특징으로 하는 API 호출 흐름 기반의 악성코드 탐지 방법.

청구항 7

제5항에 있어서,

상기 수집하는 단계는

상기 분석 대상 코드가 동작하는 동안 프로세스 별로 생성되거나 또는 기록된 상기 행위 데이터를 시간 순으로 수집하는 것

을 특징으로 하는 API 호출 흐름 기반의 악성코드 탐지 방법.

청구항 8

복수의 분석 대상 코드들이 프로그램 상에서 동작 시 호출하는 API (Application Programming Interface)의 호출 흐름에 따른 행위 데이터를 수집하는 단계;

상기 수집된 상기 복수의 분석 대상 코드들 각각의 행위 데이터에서 공통으로 호출되는 API를 추출하는 단계;

상기 추출된 공통된 API에 대응하는 제1 노드를 생성하는 단계;

상기 추출된 공통된 API의 다음에 상기 복수의 분석 대상 코드들 각각에 의하여 호출되는 API에 대응하는 적어도 하나의 제2 노드를 생성하는 단계; 및

상기 제1 노드와 상기 적어도 하나의 제2 노드를 트리 구조로 연결하여 악성코드 탐지를 위한 트리 구조의 행위 패턴을 생성하는 단계;

를 포함하는 악성코드 탐지를 위한 행위 패턴 생성 방법.

청구항 9

제8항에 있어서,

악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치를 부여하고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치를 부여하는 단계

를 더 포함하는 악성코드 탐지를 위한 행위 패턴 생성 방법.

청구항 10

제1항 내지 제9항 중 어느 한 항의 방법을 실행하기 위한 프로그램이 기록되어 있는 것을 특징으로 하는 컴퓨터에서 판독 가능한 기록매체.

청구항 11

프로그램 상의 보호된 영역 내에서 분석 대상 코드가 동작하는 동안 상기 분석 대상 코드가 호출하는 API(Application Programming Interface)의 일련의 호출 흐름에 따른 상기 분석 대상 코드의 행위 데이터를 수집하는 행위 데이터 수집부;

상기 수집된 행위 데이터를, 각 노드가 특정 행위에 대응하도록 형성된 트리 구조의 행위 패턴과 순차적으로 매칭하는 매칭부;

상기 수집된 행위 데이터와 순차적으로 매칭된 상기 행위 패턴 내의 각 노드에 부여된 가중치를 합산하는 가중치 합산부; 및

상기 합산된 가중치가 기 정해진 합산 값 이상일 경우, 상기 분석 대상 코드를 악성코드로 판단하는 악성코드 판단부;

를 포함하는 API 호출 흐름 기반의 악성코드 탐지 시스템.

청구항 12

제11항에 있어서,

상기 매칭부는

상기 트리 구조의 상기 행위 패턴 내의 현재 노드와 연결되는 다음 노드들 중, 상기 수집된 행위 데이터의 행위가 상기 현재 노드와 일치하는지의 여부에 따라 상기 다음 노드들과 순차적으로 매칭하는 것

을 특징으로 하는 API 호출 흐름 기반의 악성코드 탐지 시스템.

청구항 13

제11항에 있어서,

상기 가중치 합산부는

악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치가 부여되고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치가 부여되는 상기 트리 구조의 상기 행위 패턴을 기반으로, 상기 행위 데이터와 매칭된 상기 행위 패턴 내의 상기 가중치를 합산하는 것

을 특징으로 하는 API 호출 흐름 기반의 악성코드 탐지 시스템.

청구항 14

제11항에 있어서,

상기 행위 데이터 수집부는

상기 분석 대상 코드가 동작하는 동안 프로세스 별로 생성되거나 또는 기록된 상기 행위 데이터를 시간 순으로 수집하는 것

을 특징으로 하는 API 호출 흐름 기반의 악성코드 탐지 시스템.

청구항 15

프로그램 상의 보호된 영역 내에서 분석 대상 코드가 동작하는 동안 상기 분석 대상 코드가 호출하는 API(Application Programming Interface)를 식별하고, 상기 식별된 API에 따라서 상기 분석 대상 코드의 행위 데이터를 수집하는 행위 데이터 수집부;

각 노드가 특정 행위에 대응하도록 형성된 트리 구조의 행위 패턴 내의 현재 노드와 연결되는 다음 노드들 중, 상기 수집된 상기 분석 대상 코드의 행위 데이터와 일치하는 행위에 대응하는 노드를 식별하여 매칭하는 매칭부;

상기 매칭된 노드에 부여된 가중치를 누적하여 합산하는 가중치 합산부; 및

상기 합산된 가중치가 기 정해진 합산 값 이상일 경우, 상기 분석 대상 코드를 악성코드로 판단하는 악성코드

판단부;

를 포함하는 API 호출 흐름 기반의 악성코드 탐지 시스템.

청구항 16

제15항에 있어서,

상기 가중치 합산부는

악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치가 부여되고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치가 부여된 상기 매칭된 노드의 상기 가중치를 누적하여 합산하는 것

을 특징으로 하는 API 호출 흐름 기반의 악성코드 탐지 시스템.

청구항 17

제15항에 있어서,

상기 행위 데이터 수집부는

상기 분석 대상 코드가 동작하는 동안 프로세스 별로 생성되거나 또는 기록된 상기 행위 데이터를 시간 순으로 수집하는 것

을 특징으로 하는 API 호출 흐름 기반의 악성코드 탐지 시스템.

청구항 18

복수의 분석 대상 코드들이 프로그램 상에서 동작 시 호출하는 API (Application Programming Interface)의 호출 흐름에 따른 행위 데이터를 수집하는 행위 데이터 수집부;

상기 수집된 상기 복수의 분석 대상 코드들 각각의 행위 데이터에서 공통으로 호출되는 API를 추출하는 추출부;

상기 추출된 공통된 API에 대응하는 제1 노드를 생성하는 제1 노드 생성부;

상기 추출된 공통된 API의 다음에 상기 복수의 분석 대상 코드들 각각에 의하여 호출되는 API에 대응하는 적어도 하나의 제2 노드를 생성하는 제2 노드 생성부; 및

상기 제1 노드와 상기 적어도 하나의 제2 노드를 트리 구조로 연결하여 악성코드 탐지를 위한 트리 구조의 행위 패턴을 생성하는 행위 패턴 생성부;

를 포함하는 악성코드 탐지를 위한 행위 패턴 생성 시스템.

청구항 19

제18항에 있어서,

악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치를 부여하고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치를 부여하는 가중치 부여부;

를 더 포함하는 악성코드 탐지를 위한 행위 패턴 생성 시스템.

발명의 설명

기술 분야

[0001]

본 발명은 API 호출 흐름 기반의 악성코드 탐지 시스템 및 방법에 관한 것으로, 보다 상세하게는 프로그램 상의 보호된 영역인 샌드박스(Sandbox) 내에서 분석 대상 코드를 동작시켰을 때, API의 일련의 호출 흐름에 따른 행위 데이터를 분석하여 악성코드를 탐지하는 기술에 관한 것이다.

배경 기술

- [0002] 악성코드는 사용자가 알지 못하는 사이 컴퓨터 시스템에 침입, 설치되어 시스템이나 네트워크에 피해를 주고, 불법적으로 정보를 취득하도록 설계된 소프트웨어를 의미하며, 종래의 악성코드 또는 바이러스 탐지는 주로 파일 기반으로 수행되었다.
- [0003] 즉, 종래에는 악성코드의 탐지를 위해 알려진 모든 악성코드 파일의 패턴 또는 해쉬값 등의 특성을 추출하여 악성코드 데이터베이스에 저장해 두어야 하며, 시스템에 존재하는 모든 파일의 특성을 추출한 후 이를 악성코드 데이터베이스에 저장된 데이터와 비교하여 양자가 일치하는 경우 해당 파일을 악성코드라고 판단하였다.
- [0004] 이와 같은 종래 기술에 의하면, 악성코드 파일의 특성을 보유하고 있는 경우 해당 악성코드를 빠르고 정확하게 탐지할 수 있다는 장점이 있다. 그러나, 악성코드 파일의 특성을 보유하고 있지 않는 경우, 즉 알려지지 않은 악성코드의 경우에는 탐지 자체가 불가능하며, 기 알려진 악성코드라도 그 변종이 발생되면 동일한 유효행위를 일으키는 악성코드임에도 불구하고 탐지하기 어렵다는 단점이 있다.
- [0005] 또한, 종래 기술에 의하면, 악성코드를 탐지하기 위해 시스템에 존재하는 모든 파일에 대해 개별적으로 검사를 수행해야 하므로 악성코드 탐지 시간이 길어지는 단점이 있으며, 특히, 하루에 4천여 개 이상의 변종이 나오는 봇(Bot)과 같은 악성코드의 경우, 악성코드 탐지를 위해 모든 변종 악성코드 파일의 샘플을 보유해야 하고, 샘플 파일로부터 악성코드 탐지를 위한 파일의 특성을 일일이 추출해야 하므로, 메모리의 효율 및 탐지 효율이 떨어지는 단점이 있다.
- [0006] 한편, 한국등록특허 제10-1324691호 "모바일 악성 행위 어플리케이션 탐지 시스템 및 방법"은 사용자 단말기에서 악성 행위를 유발할 수 있는 악성 행위 어플리케이션을 탐지할 수 있는 기술로서, 모바일 악성 행위 어플리케이션의 API(Application Programming Interface) 목록 및 API 호출 순서를 패턴화하여 악성 행위 패턴을 생성하는 악성 행위 패턴 생성부 및 악성 행위 패턴에 기초하여 분석 대상 어플리케이션의 악성 행위 여부를 분석하는 악성 행위 분석부를 포함하는 기술을 제시한다.
- [0007] 상기 선행기술은 스마트폰 어플리케이션 마켓에서 유통되는 임의의 어플리케이션을 자동 수집하여 분석하고, 모바일 악성 행위 어플리케이션 여부를 확인함으로써, 모바일 악성 행위 어플리케이션 탐지 기능을 강화할 수 있는 장점이 있다. 하지만 상기 선행기술은 API 호출 순서를 패턴화하여 악성 행위 패턴을 생성한 후, 분석 대상 어플리케이션이 사용하는 API 호출 순서의 패턴이 상기 생성된 악성 행위 패턴과 일치하는지에 따라 악성코드의 여부를 판단하고 있으며, 또한 특정 API의 발생 빈도를 기반으로 악성 행위 패턴의 검출 여부를 판단하고 있기 때문에, 악성코드의 API 호출 순서와 비슷한 양성코드를 악성코드로 탐지하는 오탐율이 높은 단점이 있다.
- [0008] 따라서, 악성코드 탐지 시 오탐율은 낮추고, 정탐율은 높일 수 있는 보다 효율적인 기술 개발이 요구된다.

선행기술문헌

특허문헌

- [0009] (특허문헌 0001) 한국등록특허 제10-1324691호 (등록일: 2013.10.28)

발명의 내용

해결하려는 과제

- [0010] 본 발명은 상기와 같은 종래 기술의 문제점을 해결하고자 도출된 것으로서, API 호출 흐름 기반의 악성코드 탐지 시스템 및 방법을 제공하는 것을 목적으로 한다.
- [0011] 본 발명은 오탐율은 낮추고, 정탐율은 높일 수 있는 악성코드 탐지 기술을 제공하려는 것을 목적으로 한다.
- [0012] 본 발명은 신종 또는 변종 악성코드를 효율적으로 탐지할 수 있는 악성코드 탐지 기술을 제공하려는 것을 목적으로 한다.

과제의 해결 수단

- [0013] 이러한 목적을 달성하기 위하여 본 발명의 일 실시예에 따른 API 호출 흐름 기반의 악성코드 탐지 시스템은 프로그램 상의 보호된 영역 내에서 분석 대상 코드가 동작하는 동안 상기 분석 대상 코드가 호출하는 API(Application Programming Interface)의 일련의 호출 흐름에 따른 상기 분석 대상 코드의 행위 데이터를 수

집하는 행위 데이터 수집부, 상기 수집된 행위 데이터를, 각 노드가 특정 행위에 대응하도록 형성된 트리 구조의 행위 패턴과 순차적으로 매칭하는 매칭부, 상기 수집된 행위 데이터와 순차적으로 매칭된 상기 행위 패턴 내의 각 노드에 부여된 가중치를 합산하는 가중치 합산부 및 상기 합산된 가중치가 기 정해진 합산 값 이상일 경우, 상기 분석 대상 코드를 악성코드로 판단하는 악성코드 판단부를 포함한다.

[0014] 이때, 상기 매칭부는 상기 트리 구조의 상기 행위 패턴 내의 현재 노드와 연결되는 다음 노드들 중, 상기 수집된 행위 데이터의 행위가 상기 현재 노드와 일치하는지의 여부에 따라 상기 다음 노드들과 순차적으로 매칭할 수 있으며, 상기 가중치 합산부는 악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치가 부여되고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치가 부여되는 상기 트리 구조의 상기 행위 패턴을 기반으로, 상기 행위 데이터와 매칭된 상기 행위 패턴 내의 상기 가중치를 합산할 수 있으며, 상기 행위 데이터 수집부는 상기 분석 대상 코드가 동작하는 동안 프로세스 별로 생성되거나 또는 기록된 상기 행위 데이터를 시간 순으로 수집할 수 있다.

[0015] 또한, 본 발명의 또 다른 일 실시예에 따른 API 호출 흐름 기반의 악성코드 탐지 시스템은 프로그램 상의 보호된 영역 내에서 분석 대상 코드가 동작하는 동안 상기 분석 대상 코드가 호출하는 API(Application Programming Interface)를 식별하고, 상기 식별된 API에 따라서 상기 분석 대상 코드의 행위 데이터를 수집하는 행위 데이터 수집부, 각 노드가 특정 행위에 대응하도록 형성된 트리 구조의 행위 패턴 내의 현재 노드와 연결되는 다음 노드들 중, 상기 수집된 상기 분석 대상 코드의 행위 데이터와 일치하는 행위에 대응하는 노드를 식별하여 매칭하는 매칭부, 상기 매칭된 노드에 부여된 가중치를 누적하여 합산하는 가중치 합산부 및 상기 합산된 가중치가 기 정해진 합산 값 이상일 경우, 상기 분석 대상 코드를 악성코드로 판단하는 악성코드 판단부를 포함한다.

[0016] 이때, 상기 가중치 합산부는 악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치가 부여되고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치가 부여된 상기 매칭된 노드의 상기 가중치를 누적하여 합산할 수 있으며, 상기 행위 데이터 수집부는 상기 분석 대상 코드가 동작하는 동안 프로세스 별로 생성되거나 또는 기록된 상기 행위 데이터를 시간 순으로 수집할 수 있다.

[0017] 또한, 본 발명의 일 실시예에 따른 악성코드 탐지를 위한 행위 패턴 생성 시스템은 복수의 분석 대상 코드들이 프로그램 상에서 동작 시 호출하는 API (Application Programming Interface)의 호출 흐름에 따른 행위 데이터를 수집하는 행위 데이터 수집부, 상기 수집된 상기 복수의 분석 대상 코드들 각각의 행위 데이터에서 공통으로 호출되는 API를 추출하는 추출부, 상기 추출된 공통된 API에 대응하는 제1 노드를 생성하는 제1 노드 생성부, 상기 추출된 공통된 API의 다음에 상기 복수의 분석 대상 코드들 각각에 의하여 호출되는 API에 대응하는 적어도 하나의 제2 노드를 생성하는 제2 노드 생성부 및 상기 제1 노드와 상기 적어도 하나의 제2 노드를 트리 구조로 연결하여 악성코드 탐지를 위한 트리 구조의 행위 패턴을 생성하는 행위 패턴 생성부를 포함하며, 악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치를 부여하고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치를 부여하는 가중치 부여부를 더 포함할 수 있다.

[0018] 한편, 본 발명의 일 실시예에 따른 API 호출 흐름 기반의 악성코드 탐지 방법은 프로그램 상의 보호된 영역 내에서 분석 대상 코드를 동작시키는 단계, 상기 보호된 영역 내에서 상기 분석 대상 코드가 동작하는 동안 상기 분석 대상 코드가 호출하는 API(Application Programming Interface)의 일련의 호출 흐름에 따른 상기 분석 대상 코드의 행위 데이터를 수집하는 단계, 상기 수집된 행위 데이터를, 각 노드가 특정 행위에 대응하도록 형성된 트리 구조의 행위 패턴과 순차적으로 매칭하는 단계, 상기 수집된 행위 데이터와 순차적으로 매칭된 상기 행위 패턴 내의 각 노드에 부여된 가중치를 합산하는 단계 및 상기 합산된 가중치가 기 정해진 합산 값 이상일 경우, 상기 분석 대상 코드를 악성코드로 판단하는 단계를 포함한다.

[0019] 이때, 상기 매칭하는 단계는 상기 트리 구조의 상기 행위 패턴 내의 현재 노드와 연결되는 다음 노드들 중, 상기 수집된 행위 데이터의 행위가 상기 현재 노드와 일치하는지의 여부에 따라 상기 다음 노드들과 순차적으로 매칭할 수 있으며, 상기 합산하는 단계는 악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치가 부여되고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치가 부여되는 상기 트리 구조의 상기 행위 패턴을 기반으로, 상기 행위 데이터와 매칭된 상기 행위 패턴 내의 상기 가중치를 합산할 수 있으며, 상기 수집하는 단계는 상기 분석 대상 코드가 동작하는 동안 프로세스 별로 생성되거나 또는 기록된 상기 행위 데이터를 시간 순으로 수집할 수 있다.

[0020] 또한, 본 발명의 일 실시예에 따른 API 호출 흐름 기반의 악성코드 탐지 방법은 프로그램 상의 보호된 영역 내에서 분석 대상 코드를 동작시키는 단계, 상기 보호된 영역 내에서 상기 분석 대상 코드가 동작하는 동안 상기 분석 대상 코드가 호출하는 API(Application Programming Interface)를 식별하고, 상기 식별된 API에 따라서

상기 분석 대상 코드의 행위 데이터를 수집하는 단계, 각 노드가 특정 행위에 대응하도록 형성된 트리 구조의 행위 패턴 내의 현재 노드와 연결되는 다음 노드들 중, 상기 수집된 상기 분석 대상 코드의 행위 데이터와 일치하는 행위에 대응하는 노드를 식별하여 매칭하는 단계, 상기 매칭된 노드에 부여된 가중치를 누적하여 합산하는 단계 및 상기 합산된 가중치가 기 정해진 합산 값 이상일 경우, 상기 분석 대상 코드를 악성코드로 판단하는 단계를 포함한다.

[0021] 이때, 상기 합산하는 단계는 악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치가 부여되고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치가 부여된 상기 매칭된 노드의 상기 가중치를 누적하여 합산할 수 있으며, 상기 수집하는 단계는 상기 분석 대상 코드가 동작하는 동안 프로세스 별로 생성되거나 또는 기록된 상기 행위 데이터를 시간 순으로 수집할 수 있다.

[0022] 또한, 본 발명의 일 실시예에 따른 악성코드 탐지를 위한 행위 패턴 생성 방법은 복수의 분석 대상 코드들이 프로그램 상에서 동작 시 호출하는 API (Application Programming Interface)의 호출 흐름에 따른 행위 데이터를 수집하는 단계, 상기 수집된 상기 복수의 분석 대상 코드들 각각의 행위 데이터에서 공통으로 호출되는 API를 추출하는 단계, 상기 추출된 공통된 API에 대응하는 제1 노드를 생성하는 단계, 상기 추출된 공통된 API의 다음에 상기 복수의 분석 대상 코드들 각각에 의하여 호출되는 API에 대응하는 적어도 하나의 제2 노드를 생성하는 단계 및 상기 제1 노드와 상기 적어도 하나의 제2 노드를 트리 구조로 연결하여 악성코드 탐지를 위한 트리 구조의 행위 패턴을 생성하는 단계를 포함하며, 악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치를 부여하고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치를 부여하는 단계를 더 포함할 수 있다.

발명의 효과

[0023] 본 발명은 신종 또는 변종 악성코드를 효율적으로 탐지할 수 있는 효과가 있다.

[0024] 본 발명은 프로그램 상의 보호된 영역(sandbox) 내에서 API hooking을 통해 분석 대상 코드의 동작 흐름을 관찰하므로, 보안성이 향상된 악성코드 탐지 기술을 제공할 수 있는 효과가 있다.

[0025] 본 발명은 악성코드 탐지에 사용되는 패턴이 트리 구조의 행위 패턴으로 이루어짐에 따라, 메모리의 효율 및 탐지 효율을 향상시킬 수 있는 효과가 있다.

[0026] 본 발명은 악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치가 부여되고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치가 부여된 트리 구조의 행위 패턴을 이용해 악성코드를 탐지하므로, 오탐율을 낮추고, 정탐율은 높일 수 있는 효과가 있다.

[0027] 또한 종래 기술들은 가상의 보호된 영역인 샌드박스(sandbox) 등에서 분석 대상 코드를 실행시킨 후, 호출되는 일련의 API의 흐름을 체크하여 악성 코드 여부를 판정하는 경우에, 분석 대상 코드에 의하여 호출되는 API의 흐름을, 하나하나의 악성 코드들의 API의 흐름과 일일이 대조해야 하기 때문에 분석에 소요되는 메모리 요구량, 분석을 위한 CPU 연산, 분석 시간이 비효율적이고, 반드시 분석 대상 코드에 의하여 호출되는 API의 흐름을 끝까지 저장해 둔 후에야 하나하나의 악성 코드들의 API의 흐름과 대조할 수 있기 때문에 실시간으로 빠르게 악성코드를 탐지해 내기가 어려운 문제점이 있었다.

[0028] 이에 본 발명은 효율적으로 악성 코드 및 양성 코드의 API의 호출 패턴(행위 패턴)을 트리 구조로 저장해 두고, 분석 대상 코드에 의하여 호출되는 API의 흐름을 순차적으로 트리 구조의 행위 패턴과 매칭시켜 악성 코드 여부를 탐지하므로, 분석 시간을 매우 단축할 수 있고, 메모리 요구량과 CPU 연산량을 줄일 수 있으며, 분석 대상 코드에 의하여 호출되는 API의 흐름을 반드시 끝까지 모니터링할 필요 없이도 악성 코드인지 양성 코드인지 여부를 쉽게 탐지할 수 있다.

[0029] 또 종래 기술들은 악성 코드가 호출하는 API의 흐름 위주로 분석 대상 코드가 호출하는 API의 흐름을 비교 분석하기 때문에, 악성 코드와 호출하는 API의 흐름이 유사한 일부 양성 코드조차도 악성 코드로 판정하는 오탐 비율이 높아지는데 비하여, 본 발명은 트리 구조의 행위 패턴을 이용하여 분석 대상 코드의 API 호출 흐름을 분석하므로, 한 번의 시퀀스에서 악성 코드의 행위 패턴과 양성 코드의 행위 패턴을 한꺼번에 비교 분석해 냄으로써 양성 코드를 양성 코드로 정확히 인식하는 비율을 크게 높이고, 오탐 비율을 크게 낮출 수 있다.

도면의 간단한 설명

[0030] 도 1은 본 발명의 일 실시예에 따른 API 호출 흐름 기반의 악성코드 탐지 시스템의 개략적인 구성을 나타낸 도

면이다.

도 2는 본 발명의 일 실시예에 따른 악성코드 탐지를 위한 행위 패턴 생성 시스템의 개략적인 구성을 나타낸 도면이다.

도 3은 종래 샌드박스에 대한 개념을 나타낸 도면이다.

도 4는 종래 리스트 구조의 패턴을 나타낸 도면이다.

도 5는 본 발명의 일 실시예에 따른 트리 구조의 행위 패턴을 나타낸 도면이다.

도 6은 본 발명의 일 실시예에 따라 생성된 행위 패턴을 나타낸 도면이다.

도 7은 본 발명의 일 실시예에 따른 시스템 서비스 변조 악성 코드를 탐지하는 예를 나타낸 도면이다.

도 8은 본 발명의 일 실시예에 따른 시스템 서비스 변조 악성 코드의 행위 패턴을 매칭한 예를 나타낸 도면이다.

도 9는 본 발명의 일 실시예에 따른 API 호출 흐름 기반의 악성코드 탐지 방법에 관한 흐름도이다.

도 10은 본 발명의 일 실시예에 따른 악성코드 탐지를 위한 행위 패턴 생성 방법에 관한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0031] 이하, 본 발명의 바람직한 실시예를 첨부된 도면들을 참조하여 상세히 설명한다. 본 발명을 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략하기로 한다. 또한 본 발명의 실시예들을 설명함에 있어 구체적인 수치는 실시예에 불과하다.
- [0032] 본 발명은 API 호출 흐름 기반의 악성코드 탐지 시스템 및 방법에 관한 것으로, 보다 상세하게는 프로그램 상의 보호된 영역인 샌드박스(Sandbox) 내에서 분석 대상 코드를 동작시켰을 때, API의 일련의 호출 흐름에 따른 행위 데이터를 분석하여 악성코드를 탐지하는 기술에 관한 것이다.
- [0033] 도 1은 본 발명의 일 실시예에 따른 API 호출 흐름 기반의 악성코드 탐지 시스템의 개략적인 구성을 나타낸 도면이다.
- [0034] 우선, 도 1을 참조하여 본 발명의 악성코드 탐지 시스템(100)의 구성 요소에 대해 간단히 설명하기로 한다.
- [0035] 도 1을 참조하면, 본 발명의 API 호출 흐름 기반의 악성코드 탐지 시스템(100)은 행위 데이터 수집부(110), 매칭부(120), 가중치 합산부(130) 및 악성코드 판단부(140)를 포함한다.
- [0036] 행위 데이터 수집부(110)는 프로그램 상의 보호된 영역 내에서 분석 대상 코드가 동작하는 동안 상기 분석 대상 코드가 호출하는 API(Application Programming Interface)의 일련의 호출 흐름에 따른 상기 분석 대상 코드의 행위 데이터를 수집한다.
- [0037] 행위 데이터 수집부(110)는 행위 데이터를 수집할 때, 일련의 순서대로 API 호출 흐름(행태)을 수집할 수 있으며, 이를 이용해 매칭부(120)에서는 일련의 순서대로 수집된 행위 데이터를 이용해 트리 구조의 행위 패턴과 순차적으로 매칭하는 것이 가능하다.
- [0038] 또한, 본 발명의 또 다른 일 실시예에 따르면, 행위 데이터 수집부(110)는 상기와 같이 일련의 순서대로 API 호출 흐름을 수집하는 것이 아니라, 시간 순으로 호출되는 분석 대상 코드의 API 각각을 하나하나 식별하여, 상기 식별된 API에 따라 각각의 행위 데이터를 수집할 수 있고, 매칭부(120)는 분석 대상 코드가 호출하는 API 하나하나를 따라가면서 트리 구조의 행위 패턴과 매칭할 수 있다. 이에 대한 설명은 이후에 자세히 설명하기로 한다.
- [0039] 행위 데이터 수집부(110)에서 상기 프로그램 상의 보호된 영역은 샌드박스(Sandbox)를 의미하는 것으로, 상기 샌드박스는 보호된 영역 안에서 프로그램을 작동시켜 외부 요인으로부터 악영향을 미칠 수 있는 근원을 차단하는 보안 소프트웨어를 말한다. 즉, 샌드박스는 외부로부터 받은 파일을 바로 실행하지 않고 보호된 영역 안에서 실행시킴으로써, 악성코드 등과 같은 악영향적인 요소를 차단하는 것이 가능하다.
- [0040] 도 3은 종래 샌드박스에 대한 개념을 나타낸 도면이다.
- [0041] 도 3을 참조하면, 예를 들어, 인터넷 익스플로러의 웹 페이지에 접속하여 해당 페이지를 다운로드 받으면, 하드디스크 내에 리소스들이 채워지게 된다. 일반적으로 하드디스크 내에 리소스를 채우게 되는 과정은 Hard disk

(no sandbox) 그림에서 보여지는 것과 같을 수 있으며, 이는 다운받는 새로운 콘텐츠에 악성코드가 포함되어 있을 경우, 악의적으로 OS의 일부분, 또는 다른 어플리케이션의 정보를 훼손시킬 수 있는 문제가 발생할 수 있다.

[0042] 이러한 문제를 해결하기 위한 방안으로 샌드박스(Sandbox)의 개념이 등장하였으며, 이는 Hard disk (with sandbox) 그림에서 보여지는 것과 같을 수 있다. 즉, 샌드박스는 하드 디스크의 특정 영역을 샌드박스로 지정하고, 지정된 영역 내에서만 리소스를 사용하고 접근하도록 한다. 즉, 샌드박스는 OS와 같은 중요한 영역에는 영향을 주지 않도록 지정된 영역에서만 리소스를 사용하도록 하며, 더 자세하게는 외부에서 받은 프로그램을 JVM(Java Virtual Machine)이라는 보호된 영역 안에 가둔 뒤 작동시키므로, 프로그램이 폭주하거나 악성 바이러스의 침투를 막을 수 있다.

[0043] 또한, 행위 데이터 수집부(110)는 프로그램 상의 보호된 영역(Sandbox) 내에서 API 후킹(hooking)을 통해 분석 대상 코드가 동작하는 동안 상기 분석 대상 코드의 API 호출 정보를 수집할 수 있다.

[0044] 이때, 후킹(hooking)은 운영 체제나 응용 소프트웨어 등의 각종 컴퓨터 프로그램에서 소프트웨어 구성 요소 간에 발생하는 함수 호출, 메시지, 이벤트 등을 중간에서 바꾸거나 가로채는 명령, 방법, 기술이나 행위를 말하며, 이때 이러한 간섭된 함수 호출, 이벤트 또는 메시지를 처리하는 코드를 후크(hook)라고 한다.

[0045] 따라서, API 후킹을 이용하게 되면, API를 호출하기 전/후에 사용자의 후크 코드(hook code)를 실행시킬 수 있으며, API에 넘어온 파라미터 혹은 API 함수의 리턴 값을 엿보거나 조작할 수 있다. 또한 API 호출 자체를 취소시키거나 사용자 코드로 실행 흐름을 변경시킬 수 있다. 따라서, 본 발명에 따른 악성코드 탐지 시스템(100)은 API 후킹을 통해 수집된 API 호출 흐름을 이용함으로써, 보다 효율적으로 악성코드를 탐지하는 것이 가능하다.

[0046] 또한, 행위 데이터 수집부(110)는 상기 분석 대상 코드가 동작하는 동안 프로세스 별로 생성되거나 또는 기록된 행위 데이터(즉 API 호출 로깅(Logging)에 관한 정보)를 시간 순으로 수집할 수 있다. 이때 로깅(Logging)은 시스템을 작동할 때 시스템의 작동상태의 기록/보존, 이용자의 습성조사 및 시스템 동작의 분석 등을 하기 위해 작동 중의 각종 정보에 대한 기록을 만드는 것을 말하며, 행위 데이터 수집부(110)는 API 호출 흐름 정보가 기록된 로깅(Logging)으로부터 행위 데이터를 시간 순으로 수집할 수 있다.

[0047] 매칭부(120)는 행위 데이터 수집부(110)에서 수집된 행위 데이터를, 각 노드가 특정 행위에 대응하도록 형성된 트리 구조의 행위 패턴과 순차적으로 매칭한다.

[0048] 또한, 매칭부(120)는 트리 구조의 행위 패턴 내의 현재 노드와 연결되는 다음 노드들 중, 행위 데이터 수집부(110)에서 수집된 행위 데이터의 행위가 상기 현재 노드와 일치하는지의 여부에 따라 상기 다음 노드들과 순차적으로 매칭할 수 있다.

[0049] 이때, 본 발명에 따른 악성 코드 탐지 시스템(100)에서 악성 코드 탐지 시 사용되는 트리 구조의 행위 패턴에 대해 좀 더 자세히 살펴보면, 본 발명은 악성 코드를 탐지하기 위해 트리 형태의 구조로 형성된 행위(API 호출 흐름) 기반의 패턴을 이용한다.

[0050] 즉, 악성 코드를 탐지하기 위한 패턴 구조로서, 종래에는 리스트(List) 구조의 패턴을 이용하고 있는 반면, 본 발명에 따른 패턴 구조는 트리(Tree) 구조의 패턴을 이용한다. 이에 대한 설명은 도 4 및 도5를 참조하여 더 자세히 설명한다.

[0051] 도 4는 종래 리스트 구조의 패턴을 나타낸 도면이다.

[0052] 도 4를 참조하면, 종래 기술들은 악성 코드를 탐지하기 위한 패턴의 구조가 리스트 구조로 패턴화되어 있어, 가상의 보호된 영역인 샌드박스(sandbox) 등에서 분석 대상 코드를 실행시킨 후, 호출되는 일련의 API의 흐름을 체크하여 악성 코드 여부를 판정하는 경우에, 분석 대상 코드에 의하여 호출되는 API의 흐름을, 하나하나의 악성 코드들의 API의 흐름과 일일이 대조해야 하기 때문에 분석에 소요되는 메모리 요구량, 분석을 위한 CPU 연산, 분석 시간이 비효율적이고, 반드시 분석 대상 코드에 의하여 호출되는 API의 흐름을 끝까지 저장해 둔 후에야 하나하나의 악성 코드들의 API의 흐름과 대조할 수 있기 때문에 실시간으로 빠르게 악성 코드를 탐지해 내기가 어려운 문제점이 있었다.

[0053] 또한, 특정 API 호출 이후에 발생할 수 있는 행위는 경우의 수가 매우 많으나, 종래의 리스트 구조에서는 이를 탐지하기 위해서 패턴 패키지를 여러 번 탐색해야 하기 때문에, 시스템 성능의 문제와 결부되어 악성 코드 탐지 효율이 떨어지는 단점이 있었다. 또한, 종래 대부분의 악성 코드 탐지 기술들은 기 저장된 패턴의 일치 여부에 따라 악성 코드일 가능성을 높이기만 하는 구조로 되어 있어, 오탐 처리가 용이하지 않았으며, 이에 따라 악성 코드의 API 호출 순서와 비슷한 양성코드를 악성코드로 탐지하는 오탐율이 높은 단점이 있었다. 반면, 본 발명

에 따른 트리 형태의 행위 패턴 기반의 악성 코드 탐지 기술은 상기와 같은 종래의 문제점을 해결할 수 있다.

[0054] 도 5는 본 발명의 일 실시예에 따른 트리 구조의 행위 패턴을 나타낸 도면이다.

[0055] 도 5를 참조하면, 본 발명에 따른 악성 코드 탐지를 위한 패턴 구조는 트리(Tree) 구조의 행위 패턴으로 형성되어 있다. 본 발명은 효율적으로 악성 코드 및 양성 코드의 API의 호출 패턴(행위 패턴)을 트리 구조로 저장해 두고, 분석 대상 코드에 의하여 호출되는 API의 흐름을 순차적으로 트리 구조의 행위 패턴과 매칭시켜 악성 코드 여부를 탐지하므로, 분석 시간을 매우 단축할 수 있고, 메모리 요구량과 CPU 연산량을 줄일 수 있으며, 분석 대상 코드에 의하여 호출되는 API의 흐름을 반드시 끝까지 모니터링할 필요 없이도 악성 코드인지 양성 코드인지 여부를 쉽게 탐지할 수 있다.

[0056] 또 종래 기술들은 악성 코드가 호출하는 API의 흐름 위주로 분석 대상 코드가 호출하는 API의 흐름을 비교 분석하기 때문에, 악성 코드와 호출하는 API의 흐름이 유사한 일부 양성 코드조차도 악성 코드로 판정하는 오탐 비율이 높아지는데 비하여, 본 발명은 트리 구조의 행위 패턴을 이용하여 분석 대상 코드의 API 호출 흐름을 분석하므로, 한 번의 시퀀스에서 악성 코드의 행위 패턴과 양성 코드의 행위 패턴을 한꺼번에 비교 분석해 넘으로써 양성 코드를 양성 코드로 정확히 인식하는 비율을 크게 높이고, 오탐 비율을 크게 낮출 수 있다. 한편, 본 발명의 일 실시예에 따른 트리 구조의 행위 패턴에 대한 특성은 이후에 더 자세히 기술하기로 한다.

[0057] 가중치 합산부(130)는 행위 데이터 수집부(110)에서 수집된 행위 데이터와 순차적으로 매칭된 상기 행위 패턴 내의 각 노드에 부여된 가중치(악성지수)를 합산한다.

[0058] 또한, 가중치 합산부(130)는 악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치가 부여되고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치가 부여되는 상기 트리 구조의 상기 행위 패턴을 기반으로, 행위 데이터 수집부(110)에서 수집된 행위 데이터와 매칭된 상기 행위 패턴 내의 상기 가중치를 합산할 수 있다. 이에 대한 설명은 이후에 더 자세히 설명하기로 한다.

[0059] 악성코드 판단부(140)는 가중치 합산부(130)에서 합산된 가중치가 기 정해진 합산 값 이상일 경우, 상기 분석 대상 코드를 악성코드로 판단한다.

[0060] 한편, 이하에서는 본 발명의 일 실시예에 따른 트리 구조의 행위 패턴에 대하여, 상기 행위 패턴이 생성되는 과정 및 상기 행위 패턴의 특성에 대해 자세히 기술한다.

[0061] 우선, 도 2를 참조하여 본 발명에 따른 트리 구조의 행위 패턴을 생성하기 위한 구성 요소에 대해 살펴본다.

[0062] 도 2는 본 발명의 일 실시예에 따른 악성코드 탐지를 위한 행위 패턴 생성 시스템의 개략적인 구성을 나타낸 도면이다.

[0063] 도 2를 참조하면, 본 발명의 API 호출 흐름 기반의 악성코드 탐지 시스템(100)은 악성코드 탐지를 위한 행위 패턴 생성 시스템(200)을 포함할 수 있으며, 악성코드 탐지를 위한 행위 패턴 생성 시스템(200)은 행위 데이터 수집부(210), 추출부(220), 제1 노드 생성부(230), 제2 노드 생성부(240), 행위 패턴 생성부(250) 및 가중치 부여부(260)를 포함한다.

[0064] 행위 데이터 수집부(210)는 복수의 분석 대상 코드들이 프로그램 상에서 동작 시 호출하는 API (Application Programming Interface)의 호출 흐름에 따른 행위 데이터를 수집한다.

[0065] 즉, 행위 데이터 수집부(210)는 복수의 악성 코드 및 복수의 양성 코드를 포함하는 복수의 분석 대상 코드들이 프로그램 상에서 동작할 때, 상기 분석 대상 코드들 각각이 호출하는 API의 호출 흐름에 따른 행위 데이터를 모두 수집한다.

[0066] 추출부(220)는 행위 데이터 수집부(210)에서 수집된 상기 복수의 분석 대상 코드들 각각의 행위 데이터에서 공통으로 호출되는 API를 추출한다.

[0067] 제1 노드 생성부(230)는 추출부(220)에서 추출된 공통된 API에 대응하는 제1 노드를 생성한다.

[0068] 제2 노드 생성부(240)는 추출부(220)에서 추출된 공통된 API의 다음에 상기 복수의 분석 대상 코드들 각각에 의하여 호출되는 API에 대응하는 적어도 하나의 제2 노드를 생성한다.

[0069] 행위 패턴 생성부(250)는 제1 노드 생성부(230)에서 생성된 제1 노드와 제2 노드 생성부(240)에서 생성된 적어도 하나의 제2 노드를 트리 구조로 연결하여 악성코드 탐지를 위한 트리 구조의 행위 패턴을 생성한다. 이때, 행위 패턴 생성부(250)를 통해 생성된 행위 패턴은 도 5와 같은 트리 구조로 나타낼 수 있다.

- [0070] 가중치 부여부(260)는 악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치를 부여하고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치를 부여한다. 이때, 상기 가중치는 악성 지수를 나타낸다.
- [0071] 가중치 부여부(260)는 행위 패턴 생성부(250)에서 트리 구조의 행위 패턴이 생성된 후 가중치를 부여할 수도 있고, 또는 제1 노드 및 제2 노드를 생성할 때 가중치를 부여할 수도 있다.
- [0072] 더 자세하게는, 본 발명의 일 실시예에 따라 생성된 트리 구조의 행위 패턴 내의 각 노드는 4가지 종류로 구분할 수 있으며, 이는 플러스 패턴 노드, 마이너스 패턴 노드, 독립 패턴 노드 및 분석 패턴 노드일 수 있다.
- [0073] 상기 플러스 패턴 노드는 상/하위 패턴을 가지고 있으면서, 패턴과 일치하는 경우 악성 행위 또는 악성 행위를 위한 준비 단계로 간주되는 경우에 해당될 수 있다. 또한, 플러스 패턴 노드는 악성 코드의 행위로 간주되는 행위에 대응하는 노드로서, 플러스(+) 가중치가 부여된다.
- [0074] 즉, 기존에 악성 코드 분석 데이터를 기반으로, 10개의 악성 코드 중 7개에서 공통적으로 특정 행위(API 호출)가 발견된 경우, 상기 특정 행위는 악성 코드일 가능성이 높다는 의미로, 상기 특정 행위에 대응하는 노드에는 100점(확실히 악성코드인 경우) 중 +80의 가중치를 부여할 수 있다. 따라서, 분석 대상 코드의 행위 데이터가 악성코드 탐지 시 80점의 가중치를 가진 플러스 패턴 노드에 대응될 때, 악성 코드 탐지 시스템(100)은 상기 분석 대상 코드의 악성 코드 가능성이 80%임을 사용자에게 제공할 수도 있고, 또는 악성코드 판단부(140)에서 분석 대상 코드를 악성코드로 판단하는 기준이 '가중치(악성지수) 합산값이 78점 이상'이라는 조건으로 설정되어 있는 경우, 80의 악성지수를 지닌 상기 분석 대상 코드는 기준치를 충족하므로 이를 악성 코드로 판단하여 사용자에게 제공할 수 있다.
- [0075] 마이너스 패턴 노드는 상/하위 패턴을 가지고 있으면서, 패턴과 일치하는 경우 정상 프로그램일 가능성이 높아지거나, 혹은 확실히 정상 파일로 간주하는 경우에 해당될 수 있다. 즉, 분석 대상 코드의 행위 데이터가 악성 코드 탐지 시 마이너스 패턴 노드에 대응된다면, 상기 분석 대상 코드가 양성 코드일 가능성은 상기 마이너스 패턴 노드에 부여된 가중치의 값만큼 낮아지게 된다.
- [0076] 또한, 마이너스 패턴 노드는 양성 코드의 행위로 간주되는 행위에 대응하는 노드로서, 마이너스(-) 가중치가 부여된다.
- [0077] 본 발명은 트리 구조의 행위 패턴을 이용해 악성코드를 탐지함으로써, 한 번의 시퀀스에서 악성 코드의 행위 패턴과 양성 코드의 행위 패턴을 한꺼번에 비교 분석하는 것이 가능하다. 종래에는 리스트 구조의 패턴 매칭을 이용해 악성코드를 판단함으로써, 악성 코드와 호출하는 API의 흐름이 유사한 일부 양성 코드조차도 악성 코드로 판정하는 오탐 비율이 높은 단점이 있었다. 반면, 본 발명은 트리 구조의 행위 패턴에 마이너스 패턴 노드를 이용함으로써 오탐 비율을 크게 낮출 수 있는 효과가 있다.
- [0078] 즉, 본 발명에 따른 마이너스 패턴 노드가 악성 코드 탐지 시 상위 노드에 부여된 악성지수의 합을 무효화하거나 차감하는 경우에 대해 살펴보면, 먼저 상위 노드에 부여된 악성지수의 합을 무효화하는 경우는 확실한 시그니처에 의해 이전 의심 행위가 정상 동작으로 판단되는 경우에 무효화될 수 있다. 즉, 만약에 분석 대상 코드가 V3 관련 레지스트리에 접근하고(악성지수 +10), 디버깅 여부를 판단하고(악성지수 +20), 가상환경 여부를 판단하고(악성지수 +5), 실행 파일을 다운로드(악성지수 +20) 하는 행위를 순차적으로 나타낸 경우, 상기의 행위로 악성지수 합산 값이 55이므로, 상기 분석 대상 코드는 악성코드일 가능성이 높다. 하지만 실행 파일을 다운로드 하는 행위 이후에 다운로드 된 실행파일에 Ahnlab 인증서가 포함되어 행위가 감지된다면, 상기 분석 대상 코드는 악성 코드가 아닌 확실한 양성 코드이므로, 상기 인증서를 감지하는 행위에는 상위 노드에 부여된 악성지수의 합을 무효화시키는 마이너스 패턴 노드가 형성될 수 있다.
- [0079] 또한, 상위 노드에 부여된 악성지수의 합을 감산하는 경우는 이전 행위로 보아 악성코드로 의심이 되지만, 후행에 뚜렷한 악성행위가 없어 수동으로 분석해야 할 대상인 경우일 수 있다. 즉, 만약에 분석 대상 코드가 서비스 관리자 해독을 획득(악성지수 +10)하고, 특정 시스템 서비스의 핸들을 획득(악성지수 +10)하고, 서비스를 중지(악성지수 +50)하는 행위를 순차적으로 나타낸 경우, 상기의 행위로 악성지수 합산 값이 70이므로, 상기 분석 대상 코드는 악성코드일 가능성이 상당히 높다. 그러나 서비스를 중지하는 행위 이후에 서비스 실행 관련 레지스트리나 파일의 변조 없이 재시작 하는 행위(악성지수 -20)가 나타난다면, 재시작 하는 행위 이후에 뚜렷한 악성행위가 없으므로, 상기 재시작 하는 행위에는 마이너스 패턴 노드가 형성될 수 있다.
- [0080] 한편, 독립 패턴 노드는 트리에서 부모 노드가 존재하지 않는 최상의 노드를 의미한다. 즉, 독립 패턴 노드는

Root 노드를 의미하며, leaf 노드를 가지고 있지 않을 수 있다.

[0081] 또한, 분석 패턴 노드는 상기 독립 패턴 노드와 유사하나, 가감할 악성지수를 가지고 있지 않은 패턴 노드를 의미하며, 또한 수동 분석 또는 리포트 생성 시 정보를 표현하기 위한 패턴 노드를 의미한다.

[0082] 도 6은 본 발명의 일 실시예에 따라 생성된 행위 패턴을 나타낸 도면이다.

[0083] 도 6을 참조하면, 본 발명의 일 실시예에 따라 생성된 트리 구조의 행위 패턴 내의 각 노드들은 특정 행위에 대응하도록 형성되어 있으며, P0 노드는 부모 노드가 존재하지 않는 독립 패턴 노드임을 알 수 있고, P1, P3, P7 노드는 마이너스 가중치가 부여된 마이너스 패턴 노드임을 알 수 있다. 또한, 그 밖에 P2, P4 내지 P6, P8 내지 P12는 플러스 가중치가 부여된 플러스 패턴 노드임을 알 수 있다.

[0084] 또한, 이때 상위 패턴이 존재하지 않는 최상위 패턴은 Root 노드(P0 노드)로 나타내고, 하위 패턴이 존재하지 않는 최하위 패턴은 Leaf 노드(P1, P5, P9, P10, P12 노드)로 나타내고, 상/하위 패턴이 모두 존재하는 패턴은 Node 패턴(P2, P3, P4, P6, P7, P8, P11 노드)으로 나타낼 수 있다.

[0085] 예를 들어, 악성코드 판단부(140)에 기 저장된 악성 코드 판단 기준의 기준치가 악성지수 합이 80 이상이 경우로 설정되어 있고, 행위 데이터 수집부(110)에서 분석 대상 코드가 호출하는 API의 일련의 호출 흐름에 따른 행위 데이터로 A,B,C의 행위 데이터를 수집했을 경우, 매칭부(120)는 상기 수집된 A,B,C의 행위 데이터를, 트리 구조의 행위 패턴과 순차적으로 매칭한다. 이때, 매칭부(120)에서 매칭된 패턴이 A 행위 데이터는 P0 노드와 매칭되고, B 행위 데이터는 P3 노드와 매칭되고, C 행위 데이터는 P6 노드에 매칭된 경우, 가중치 합산부(130)는 각각의 노드에 부여된 가중치(악성지수)를 합산한다. 즉, P0 노드는 +20이고, P3 노드는 -20이며, P6 노드는 +45이므로, 분석 대상 코드의 가중치(악성지수) 합산 값은 45가 된다. 이는 악성 코드의 판단 기준인 80 이상의 악성지수 값을 충족하지 못하므로, 악성코드 판단부(140)는 상기 분석 대상 코드를 양성 코드로 판단한다.

[0086] 상기와 같은 예는 행위 데이터 수집부(110)가 행위 데이터를 수집할 때, 일련의 순서대로 API 호출 흐름(행태)을 수집하고, 매칭부(120)에서는 일련의 순서대로 수집된 행위 데이터를 트리 구조의 행위 패턴과 순차적으로 모두 매칭한 후, 매칭된 모든 노드에 대한 가중치를 합산하여 악성 코드를 판단하는 예에 대해 설명한 것이다.

[0087] 한편, 본 발명의 또 다른 일 실시예에 따르면, 행위 데이터 수집부(110)는 상기와 같이 일련의 순서대로 API 호출 흐름을 수집하는 것이 아니라, 시간 순으로 호출되는 분석 대상 코드의 API 각각을 하나하나 식별하여, 상기 식별된 API에 따라 각각의 행위 데이터를 수집할 수 있고, 매칭부(120)는 분석 대상 코드가 호출하는 API 하나하나를 따라가면서 트리 구조의 행위 패턴과 매칭할 수 있다.

[0088] 예를 들어, 악성코드 판단부(140)에 기 저장된 악성 코드 판단 기준의 기준치가 악성지수 합이 80 이상이 경우로 설정되어 있고, 분석 대상 코드가 호출하는 API의 호출 흐름이 A, B, C, D, E, F의 순으로 호출할 경우, 행위 데이터 수집부(110)는 우선 API가 호출하는 A 행위 데이터를 수집하고, 매칭부(120)는 A 행위 데이터와 일치하는 행위에 대응하는 노드를 식별하여 매칭한다. 이때 A 행위 데이터와 매칭된 노드는 P1(+20) 노드라고 가정한다.

[0089] 다음으로 다시 행위 데이터 수집부(110)는 A 행위 데이터 다음으로 호출하는 API, 즉 B 행위 데이터를 수집하고, 매칭부(120)는 B 행위 데이터와 일치하는 행위에 대응하는 노드를 식별하여 매칭하며, 이때 B 행위 데이터와 매칭된 노드는 P2(+30) 노드라고 가정한다. 그리고 가중치 합산부(130)는 매칭부(120)에서 식별하여 매칭된 노드의 가중치를 누적하여 합산하며, 이때 합산된 가중치 값은 $P1(+20) + P2(+30) = 50$ 이 된다.

[0090] 다음으로, 행위 데이터 수집부(110)는 B 행위 데이터 다음으로 호출하는 API, 즉 C 행위 데이터를 수집하고, 매칭부(120)는 C 행위 데이터와 일치하는 행위에 대응하는 노드를 식별하여 매칭하며, 이때 C 행위 데이터와 매칭된 노드는 P3(+40) 노드라고 가정한다. 그리고 가중치 합산부(130)는 이전에 누적되어 합산된 50에 P3(+40)의 가중치 값을 더하며, 이에 따라 P1 내지 P3의 가중치의 합산 값이 90이 된다. 다음으로, 악성코드 판단부(140)는 가중치 합산부(130)에서 합산된 가중치 합산값 90 이 기 정해진 악성 코드 판단 수치(80) 이상이므로, 상기 분석 대상 코드를 악성코드로 판단한다.

[0091] 즉, 행위 데이터 수집부(110)는 분석 대상 코드가 호출하는 API가 트리 구조의 행위 패턴에 모두 매칭될 때까지 계속 행위 데이터를 수집할 수 있으며, 마찬가지로 매칭부(120)도 API가 행위 패턴에 모두 매칭될 때까지, 상기 분석 대상 코드의 행위 데이터를 트리 구조의 행위 패턴에 매칭시킬 수 있다. 단, 행위 데이터 수집부(110)는 가중치 합산부(130)에서의 가중치 합산 값이 기 정해진 악성 코드를 판단 기준 이상이 될 경우, 행위 데이터를 수집하는 행위 및 매칭하는 행위를 멈출 수 있다.

- [0092] 따라서 본 발명은 한 번의 시퀀스에서 악성 코드의 행위 패턴과 양성 코드의 행위 패턴을 한꺼번에 비교 분석해 내는 것이 가능하며, 양성 코드를 양성 코드로 정확히 인식하는 비율을 크게 높이고, 오탐 비율을 크게 낮출 수 있는 효과가 있다.
- [0093] 도 7은 본 발명의 일 실시예에 따른 시스템 서비스 변조 악성 코드를 탐지하는 예를 나타낸 도면이다.
- [0094] 도 7을 참조하면, 시스템 서비스를 변조하는 악성 코드가 샌드박스 내에서 동작할 때, 상기 악성 코드는 다음과 같이 5 단계의 패턴 양상을 보일 수 있다.
- [0095] 먼저, 시스템 서비스를 변조하는 악성 코드는 처음 호출하는 API가 Windows system directory 하위에 실행 파일을 생성하는 패턴 양상을 보일 수 있다. 상기 패턴 양상은 매칭부(120)에 의하여 악성코드 탐지를 위한 행위 패턴 내의 Root P0 노드(악성지수 +20)에 매칭될 수 있다. 다음으로 두번째로 호출하는 API는 Windows service 관리자의 요청 handle을 획득하는 패턴 양상을 보일 수 있으며, 이와 같은 패턴 양상은 매칭부(120)에 의하여 행위 패턴 내의 Node P2 노드(악성지수 +10)에 매칭될 수 있다. 다음으로 세번째로 호출하는 API는 이미 존재하는 system service의 handle을 획득하고 중지시키는 패턴 양상을 보일 수 있으며, 이와 같은 패턴 양상은 매칭부(120)에 의하여 Node P4 노드(악성지수 +20)에 매칭될 수 있다. 다음으로, 네번째로 호출하는 API는 시스템 서비스 관련 registry data 변조를 통해 상위에서 생성한 실행 파일을 시스템 서비스로 은폐하는 패턴 양상을 보일 수 있으며, 이와 같은 패턴 양상은 매칭부(120)에 의하여 Node P8 노드(악성지수 +40)에 매칭될 수 있다. 다음으로, 다섯번째로 호출하는 API는 변조된 서비스를 재시작 하는 패턴 양상을 보일 수 있으며, 이와 같은 패턴 양상은 Leaf P12 노드(악성지수 +10)에 매칭될 수 있다. 이때, 각각의 패턴 양상을 트리 구조의 행위 패턴에 매칭한 예는 도 8과 같을 수 있다.
- [0096] 도 8은 본 발명의 일 실시예에 따른 시스템 서비스 변조 악성 코드의 행위 패턴을 매칭한 예를 나타낸 도면이다.
- [0097] 따라서, 도 7을 참조하면, 시스템 서비스를 변조하는 악성 코드는 트리 구조의 행위 패턴 내에 P0(+20), P2(+10), P4(+20), P8(+40) 및 P12(+10) 노드에 매칭되는 것을 알 수 있고, 상기 시스템 서비스를 변조하는 악성 코드의 악성 지수(가중치)는 100 임을 알 수 있다. 따라서, 악성코드 판단부(140)에 악성 코드 판단 기준으로서 가중치 합산 값이 100 이상일 경우로 기 정해져 있을 경우, 악성코드 판단부(140)는 상기 시스템 서비스를 변조하는 악성 코드를 악성 코드로 판단한다.
- [0098] 이하에서는 악성 코드 탐지 시스템(100)이 악성코드를 탐지할 때, 행위 데이터 수집부(110)가 일련의 순서대로 API 호출 흐름을 수집하는 것이 아니라, 시간 순으로 호출되는 분석 대상 코드의 API 각각을 하나하나 식별하여, 상기 식별된 API에 따라 각각의 행위 데이터를 수집하고, 매칭부(120)가 분석 대상 코드가 호출하는 API 하나하나를 따라가면서 트리 구조의 행위 패턴과 매칭하는 경우에 대하여, 각 구성 요소별 역할을 간단히 설명하기로 하며, 이에 대한 실시예는 상기에 설명했으므로 이를 참조하도록 한다.
- [0099] 즉, 본 발명의 또 다른 일 실시예에 따른 API 호출 흐름 기반의 악성코드 탐지 시스템(100)은 행위 데이터 수집부(110), 매칭부(120), 가중치 합산부(130) 및 악성코드 판단부(140)를 포함한다.
- [0100] 행위 데이터 수집부(110)는 프로그램 상의 보호된 영역 내에서 분석 대상 코드가 동작하는 동안 상기 분석 대상 코드가 호출하는 API(Application Programming Interface)를 식별하고, 상기 식별된 API에 따라서 상기 분석 대상 코드의 행위 데이터를 수집한다.
- [0101] 또한, 행위 데이터 수집부(110)는 분석 대상 코드가 동작하는 동안 프로세스 별로 생성되거나 또는 기록된 상기 행위 데이터를 시간 순으로 수집할 수 있다.
- [0102] 매칭부(120)는 각 노드가 특정 행위에 대응하도록 형성된 트리 구조의 행위 패턴 내의 현재 노드와 연결되는 다음 노드들 중, 상기 수집된 상기 분석 대상 코드의 행위 데이터와 일치하는 행위에 대응하는 노드를 식별하여 매칭한다.
- [0103] 가중치 합산부(130)는 매칭부(120)에서 매칭된 노드에 부여된 가중치(악성지수)를 누적하여 합산한다.
- [0104] 또한, 가중치 합산부(130)는 악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치가 부여되고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치가 부여된 상기 매칭된 노드의 상기 가중치를 누적하여 합산할 수 있다.
- [0105] 악성코드 판단부(140)는 가중치 합산부(130)에서 합산된 가중치가 기 정해진 합산 값 이상일 경우, 상기 분석

대상 코드를 악성코드로 판단한다.

- [0106] 이하에서는 본 발명의 일 실시예에 따른 API 호출 흐름 기반의 악성코드 탐지 방법에 대해 설명하되, 상기에 자세히 기술한 내용을 바탕으로 간단히 설명하기로 한다.
- [0107] 도 9는 본 발명의 일 실시예에 따른 API 호출 흐름 기반의 악성코드 탐지 방법에 관한 흐름도이다.
- [0108] 도 9를 참조하면, 본 발명의 일 실시예에 따른 API 호출 흐름 기반의 악성코드 탐지 방법은 먼저 행위 데이터 수집부(110)에서 프로그램 상의 보호된 영역 내에서 분석 대상 코드를 동작시킨다(S910).
- [0109] 이때, 상기 프로그램 상의 보호된 영역은 샌드박스(Sandbox)를 의미하는 것으로, 상기 샌드박스는 보호된 영역 안에서 프로그램을 작동시켜 외부 요인으로부터 악영향을 미칠 수 있는 근원을 차단하는 보안 소프트웨어를 말한다. 즉, 샌드박스는 외부로부터 받은 파일을 바로 실행하지 않고 보호된 영역 안에서 실행시킴으로써, 악성코드 등과 같은 악영향적인 요소를 차단하는 것이 가능하며, 이에 대한 설명은 상기에 도3을 참조하여 자세히 설명했으므로, 이를 참조하기로 한다.
- [0110] 다음으로, 행위 데이터 수집부(110)는 상기 보호된 영역 내에서 상기 분석 대상 코드가 동작하는 동안 상기 분석 대상 코드가 호출하는 API(Application Programming Interface)의 일련의 호출 흐름에 따른 상기 분석 대상 코드의 행위 데이터를 수집한다(S920).
- [0111] 이때, 행위 데이터 수집부(110)는 행위 데이터를 수집할 때, 일련의 순서대로 API 호출 흐름(행태)을 수집할 수 있으며, 이를 이용해 매칭부(120)에서는 일련의 순서대로 수집된 행위 데이터를 이용해 트리 구조의 행위 패턴과 순차적으로 매칭하는 것이 가능하다.
- [0112] 또한 본 발명의 또 다른 일 실시예에 따르면, 행위 데이터 수집부(110)는 상기와 같이 일련의 순서대로 API 호출 흐름을 수집하는 것이 아니라, 시간 순으로 호출되는 분석 대상 코드의 API 각각을 하나하나 식별하여, 상기 식별된 API에 따라 각각의 행위 데이터를 수집할 수 있고, 매칭부(120)는 분석 대상 코드가 호출하는 API 하나하나를 따라가면서 트리 구조의 행위 패턴과 매칭할 수 있다.
- [0113] 즉, 행위 데이터 수집부(110)는 프로그램 상의 보호된 영역 내에서 분석 대상 코드가 동작하는 동안 상기 분석 대상 코드가 호출하는 API(Application Programming Interface)를 식별하고, 상기 식별된 API에 따라서 상기 분석 대상 코드의 행위 데이터를 수집할 수 있다. 이에 대한 설명은 상기에 자세히 설명했으므로, 이하 생략하기로 한다.
- [0114] 또한, 이때 행위 데이터 수집부(110)는 프로그램 상의 보호된 영역(Sandbox) 내에서 API 후킹(hooking)을 통해 분석 대상 코드가 동작하는 동안 상기 분석 대상 코드의 API 호출 정보를 수집할 수 있으며, 행위 데이터 수집부(110)는 분석 대상 코드가 동작하는 동안 프로세스 별로 생성되거나 또는 기록된 상기 행위 데이터를 시간 순으로 수집할 수 있다.
- [0115] 다음으로, 매칭부(120)는 상기 수집된 행위 데이터를, 각 노드가 특정 행위에 대응하도록 형성된 트리 구조의 행위 패턴과 순차적으로 매칭한다(S930).
- [0116] 이때 매칭부(120)는 트리 구조의 행위 패턴 내의 현재 노드와 연결되는 다음 노드들 중, 행위 데이터 수집부(110)에서 수집된 행위 데이터의 행위가 상기 현재 노드와 일치하는지의 여부에 따라 상기 다음 노드들과 순차적으로 매칭할 수 있다.
- [0117] 또한 본 발명의 또 다른 일 실시예에 따르면, 매칭부(120)는 각 노드가 특정 행위에 대응하도록 형성된 트리 구조의 행위 패턴 내의 현재 노드와 연결되는 다음 노드들 중, 상기 수집된 상기 분석 대상 코드의 행위 데이터와 일치하는 행위에 대응하는 노드를 식별하여 매칭할 수 있다.
- [0118] 다음으로, 가중치 합산부(130)는 행위 데이터 수집부(110)에서 수집된 행위 데이터와 순차적으로 매칭된 상기 행위 패턴 내의 각 노드에 부여된 가중치를 합산한다(S940).
- [0119] 또한, 가중치 합산부(130)는 악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치가 부여되고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치가 부여되는 상기 트리 구조의 상기 행위 패턴을 기반으로, 행위 데이터 수집부(110)에서 수집된 행위 데이터와 매칭된 상기 행위 패턴 내의 상기 가중치를 합산할 수 있다.
- [0120] 또한, 가중치 합산부(130)는 행위 데이터 수집부(110)가 프로그램 상의 보호된 영역 내에서 분석 대상 코드가

동작하는 동안 상기 분석 대상 코드가 호출하는 API(Application Programming Interface)를 식별하고, 상기 식별된 API에 따라서 상기 분석 대상 코드의 행위 데이터를 수집할 경우, 매칭된 노드에 부여된 가중치를 누적하여 합산할 수 있다.

- [0121] 다음으로, 악성코드 판단부(140)는 상기 합산된 가중치가 기 정해진 합산 값 이상일 경우, 상기 분석 대상 코드를 악성코드로 판단한다(S950).
- [0122] 본 발명은 한 번의 시퀀스에서 악성 코드의 행위 패턴과 양성 코드의 행위 패턴을 한꺼번에 비교 분석해 내는 것이 가능하며, 양성 코드를 양성 코드로 정확히 인식하는 비율을 크게 높이고, 오탐 비율을 크게 낮출 수 있는 효과가 있다.
- [0123] 도 10은 본 발명의 일 실시예에 다른 악성코드 탐지를 위한 행위 패턴 생성 방법에 관한 흐름도이다.
- [0124] 도 10을 참조하면, 본 발명의 일 실시예에 다른 악성코드 탐지를 위한 행위 패턴 생성 방법은 상기 API 호출 흐름 기반의 악성코드 탐지 방법에 포함될 수 있으며, 먼저 복수의 분석 대상 코드들이 프로그램 상에서 동작 시 호출하는 API (Application Programming Interface)의 호출 흐름에 따른 행위 데이터를 수집한다(S1010).
- [0125] 이때, 행위 데이터 수집부(210)는 복수의 악성 코드 및 복수의 양성 코드를 포함하는 복수의 분석 대상 코드들이 프로그램 상에서 동작할 때, 상기 분석 대상 코드들 각각이 호출하는 API의 호출 흐름에 따른 행위 데이터를 모두 수집한다.
- [0126] 다음으로, 추출부(220)는 상기 수집된 상기 복수의 분석 대상 코드들 각각의 행위 데이터에서 공통으로 호출되는 API를 추출(S1020)하고, 다음으로, 제1 노드 생성부(230)는 상기 추출된 공통된 API에 대응하는 제1 노드를 생성(S1030)하며, 다음으로, 제2 노드 생성부(240)는 상기 추출된 공통된 API의 다음에 상기 복수의 분석 대상 코드들 각각에 의하여 호출되는 API에 대응하는 적어도 하나의 제2 노드를 생성(S1040)한다.
- [0127] 다음으로, 행위 패턴 생성부(250)는 제1 노드 생성부(230)에서 생성된 제1 노드와 제2 노드 생성부(240)에서 생성된 적어도 하나의 제2 노드를 트리 구조로 연결하여 악성코드 탐지를 위한 트리 구조의 행위 패턴을 생성한다(S1050).
- [0128] 이때, 행위 생성 패턴부(250)를 통해 생성된 행위 패턴은 도 5와 같은 트리 구조로 나타낼 수 있으며, 본 발명에 따라 생성된 트리 구조의 행위 패턴은 플러스 패턴 노드, 마이너스 패턴 노드, 독립 패턴 노드 및 분석 패턴 노드와 같이 4가지 종류로 구분될 수 있다. 이에 대한 설명은 상기에 자세히 설명했으므로 이하 생략하기로 한다.
- [0129] 다음으로, 가중치 부여부(260)는 악성 코드의 행위로 간주되는 행위에 대응하는 노드에는 플러스(+) 가중치를 부여하고, 양성 코드의 행위로 간주되는 행위에 대응하는 노드에는 마이너스(-) 가중치를 부여한다(S1060).
- [0130] 이때, 상기 가중치는 악성 지수를 나타내며, 가중치 부여부(260)는 행위 패턴 생성부(250)에서 트리 구조의 행위 패턴이 생성된 후 가중치를 부여할 수도 있고, 또는 제1 노드 및 제2 노드를 생성할 때 가중치를 부여할 수도 있다.
- [0131] 본 발명의 일 실시 예에 따른 API 호출 흐름 기반의 악성코드 탐지 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 본 발명의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0132] 이상과 같이 본 발명에서는 구체적인 구성 요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및

변형이 가능하다.

[0133]

따라서, 본 발명의 사상은 설명된 실시예에 국한되어 정해져서는 아니 되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 것들은 본 발명 사상의 범주에 속한다고 할 것이다.

부호의 설명

[0134]

100: API 호출 흐름 기반의 악성코드 탐지 시스템

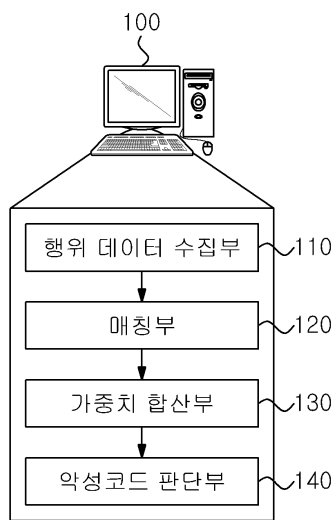
110: 행위 데이터 수집부 120: 매칭부

130: 가중치 합산부

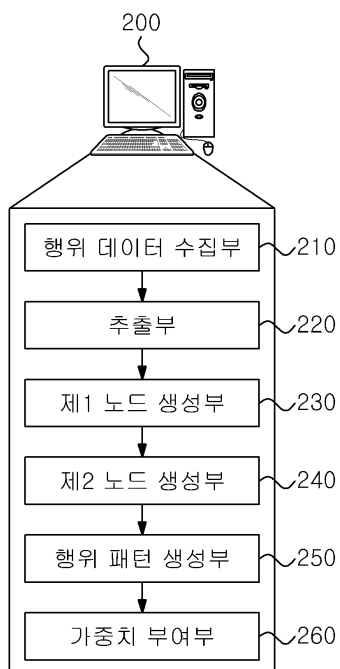
140: 악성코드 판단부

도면

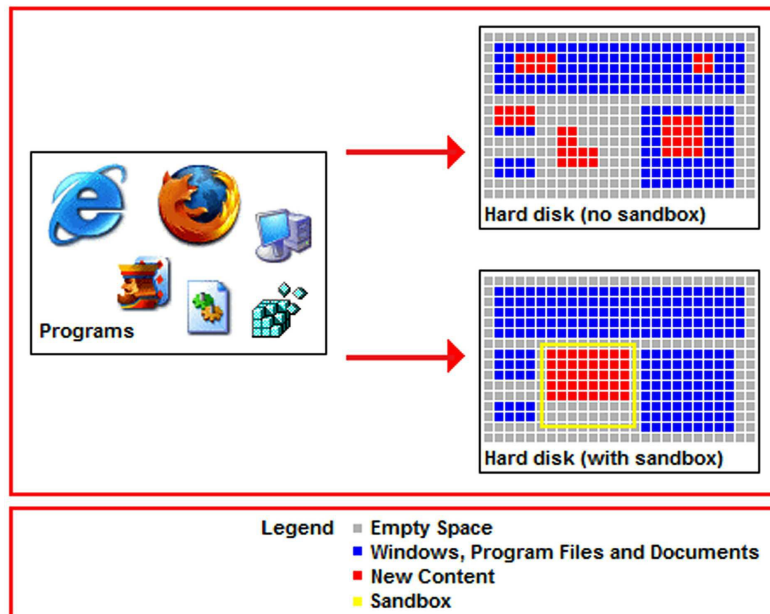
도면1



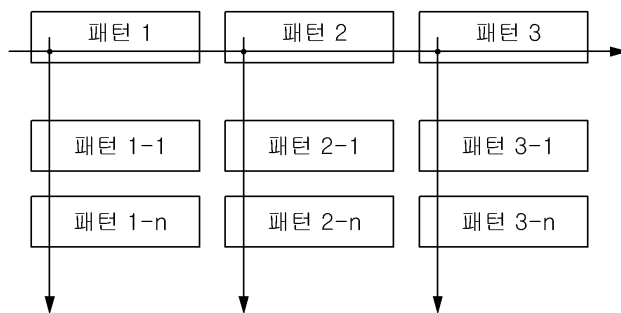
도면2



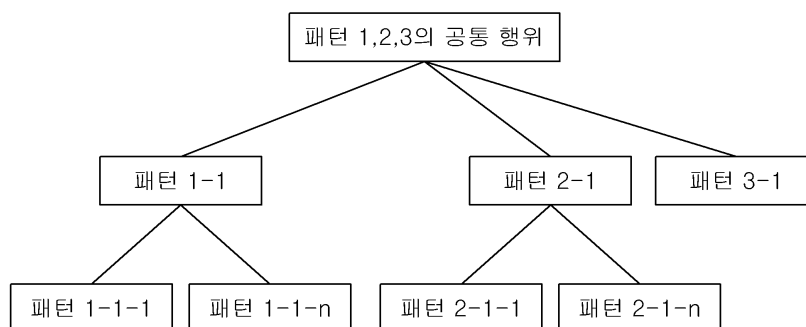
도면3



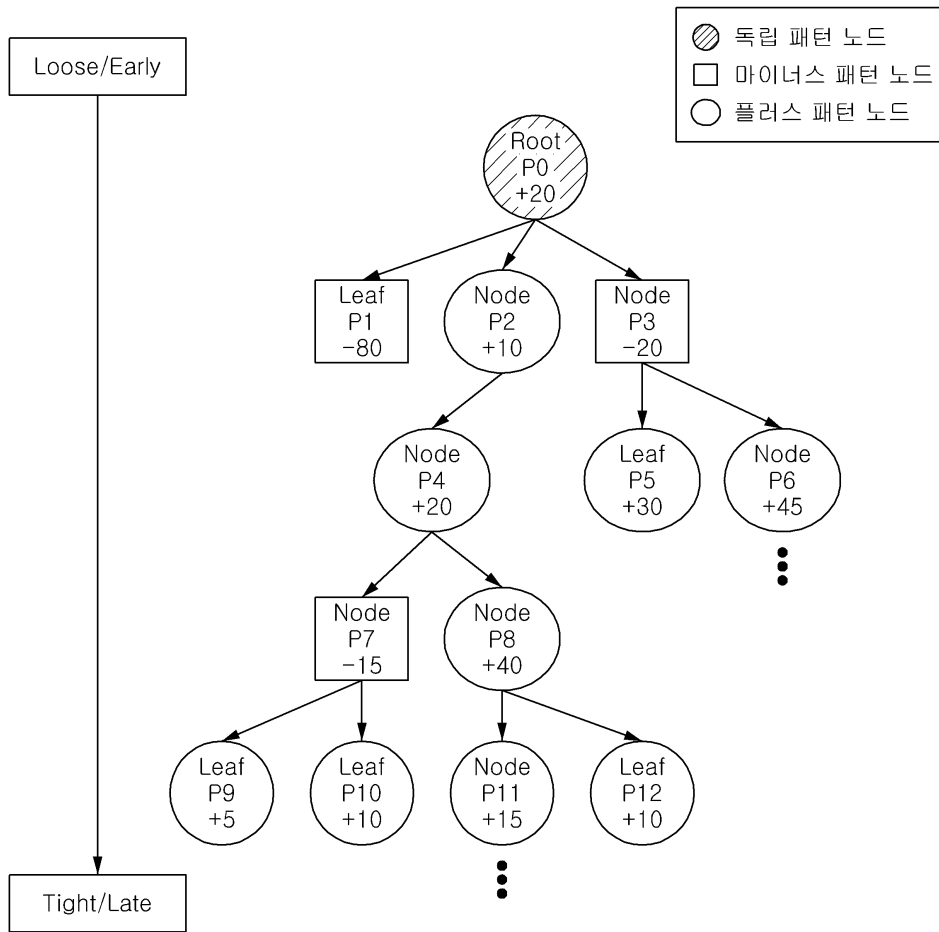
도면4



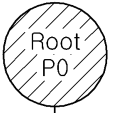
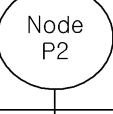
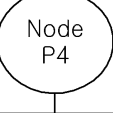
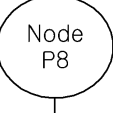
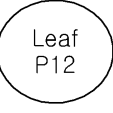
도면5



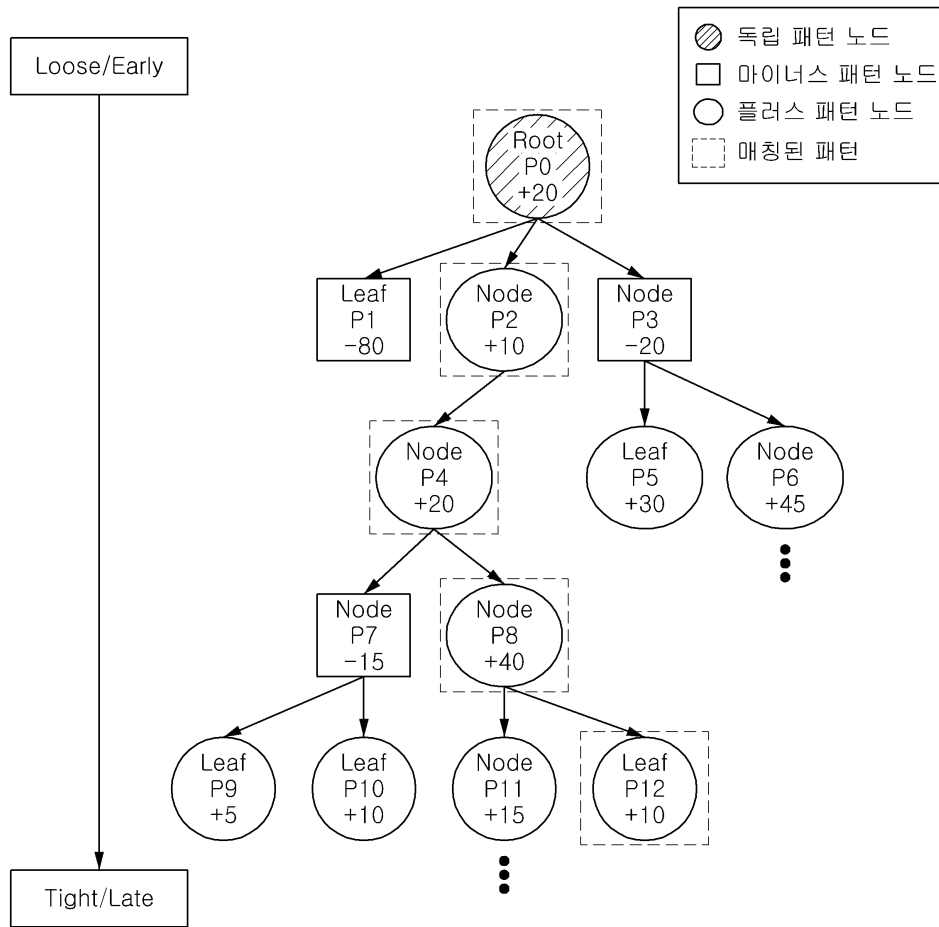
도면6



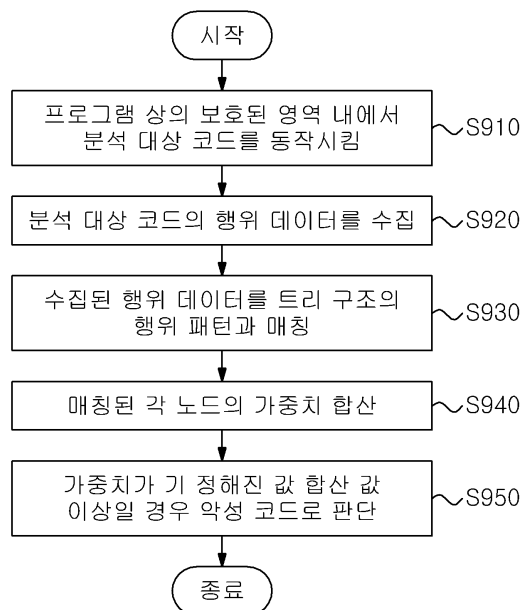
도면7

Node	패턴	악성지수
	Windows system directory 하위에 실행 파일을 생성	+20
	Windows service 관리자 요청 handle 획득	+10
	이미 존재하는 system service의 handle을 획득하고 중지 시킴	+20
	시스템 서비스관련 registry data 변조를 통해 상위에서 생성한 실행 파일을 시스템 서비스로 은폐	+40
	변조 된 서비스 재 시작	+10

도면8



도면9



도면10

