(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년10월12일
(11) 등록번호 10-1907681
(24) 등록일자 2018년10월05일

(51) 국제특허분류(Int. Cl.)
G06F 21/56 (2013.01) *G06F 17/30* (2006.01)
(52) CPC특허분류
G06F 21/56 (2013.01)
G06F 17/30705 (2013.01)
(21) 출원번호 10-2017-0024455
(22) 출원일자 2017년02월24일
심사청구일자 2017년02월24일
(65) 공개번호 10-2018-0097824
(43) 공개일자 2018년09월03일
(56) 선행기술조사문헌
KR1020070049511 A
KR1020160031588 A

(73) 특허권자
성균관대학교산학협력단
경기도 수원시 장안구 서부로 2066 (천천동, 성균관대학교내)
(72) 발명자
조금환
경기도 수원시 장안구 서부로 2066 천천동 300 성균관대학교자연과학캠퍼스 27317호
조준성
경기도 수원시 장안구 서부로 2066 천천동 300 성균관대학교자연과학캠퍼스 27317호
김형식
경기도 수원시 장안구 서부로 2066 천천동 300 성균관대학교자연과학캠퍼스 27324호
(74) 대리인
특허법인로얄

전체 청구항 수 : 총 17 항 심사관 : 홍경아

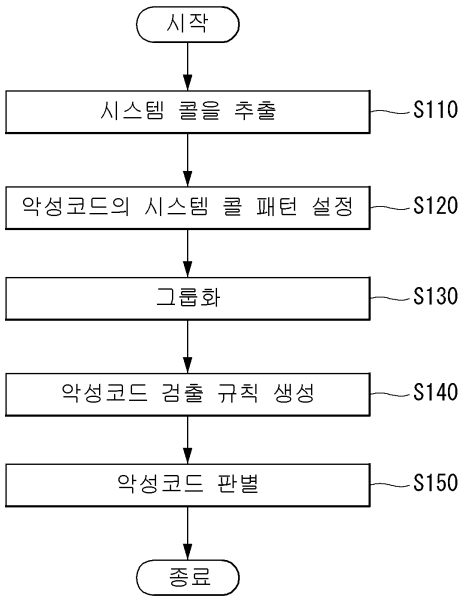
(54) 발명의 명칭 **악성코드 검출을 위한 자동 규칙 생성방법, 장치, 시스템 및 이를 기록한 컴퓨터로 판독가능한 기록매체**

(57) 요 약

본 발명은 악성코드 검출을 위한 자동 규칙 생성방법, 장치, 시스템 및 이를 기록한 컴퓨터로 판독가능한 기록매체을 개시한다. 본 발명에 악성코드 검출을 위한 자동 규칙 생성방법은 따른 다수의 악성코드와 정상 프로그램을 실행시켜 다수의 시스템 콜을 추출하는 단계, 상기 추출된 다수의 시스템 콜 각각이 상기 다수의 악성코드와 상

(뒷면에 계속)

대 표 도 - 도1



기 정상 프로그램에서 실행된 횟수를 이용하여 점수를 부여하는 단계, 상기 부여된 점수 차이가 임계값 이상인 시스템 콜을 추출하고, 상기 다수의 악성코드 중 각각의 악성코드의 추출된 시스템 콜의 패턴을 상기 각각의 악성코드의 시스템 콜 패턴으로 설정하는 단계, 상기 각각의 악성코드의 시스템 콜 패턴을 유사도를 기반으로 그룹화하는 단계, 상기 그룹화된 악성코드의 시스템 콜 패턴을 다중서열정렬(Multiple sequence alignment) 알고리즘을 이용하여 각 그룹의 대표 패턴을 선정하여 악성코드 검출 규칙으로 생성하는 단계, 및 상기 생성된 악성코드 검출 규칙과 새로 실행되는 프로그램의 시스템 콜 패턴을 비교하여 악성코드를 판별하는 단계를 포함한다. 본 발명에 따르면, 시스템 콜의 패턴을 이용하여 악성코드를 검출함으로써 악성코드에 사용되지 않던 시스템 콜이 중간에 삽입되는 경우에도 패턴을 이용하여 악성코드를 검출할 수 있다.

이 발명을 지원한 국가연구개발사업

과제고유번호	IITP-2016-R0092-16-1006
부처명	미래창조과학부
연구관리전문기관	정보통신기술진흥센터
연구사업명	대학ICT연구센터육성지원사업
연구과제명	사물인터넷 보안기술 연구
기 여 율	1/1
주관기관	순천향대학교 산학협력단
연구기간	2015.06.01 ~ 2018.12.31

공지예외적용 : 있음

명 세 서

청구범위

청구항 1

다수의 악성코드와 정상 프로그램을 실행시켜 다수의 시스템 콜을 추출하는 단계;

상기 추출된 다수의 시스템 콜 각각이 상기 다수의 악성코드와 상기 정상 프로그램에서 실행된 횟수를 이용하여 점수를 부여하는 단계;

상기 부여된 점수 차이가 임계값 이상인 시스템 콜을 추출하고, 상기 다수의 악성코드 중 각각의 악성코드의 추출된 시스템 콜의 패턴을 상기 각각의 악성코드의 시스템 콜 패턴으로 설정하는 단계;

상기 각각의 악성코드의 시스템 콜 패턴을 유사도를 기반으로 그룹화하는 단계;

상기 그룹화된 악성코드의 시스템 콜 패턴을 다중서열정렬(Multiple sequence alignment) 알고리즘을 이용하여 각 그룹의 대표 패턴을 선정하여 악성코드 검출 규칙으로 생성하는 단계; 및

상기 생성된 악성코드 검출 규칙과 새로 실행되는 프로그램의 시스템 콜 패턴을 비교하여 악성코드를 판별하는 단계;

를 포함하는 악성코드 검출을 위한 자동 규칙 생성방법.

청구항 2

제1항에 있어서,

상기 악성코드를 판별하는 단계 다음에,

상기 새로 실행되는 프로그램이 악성코드인 것으로 판별된 경우, 상기 새로 실행되는 프로그램의 시스템 콜 패턴을 상기 악성코드의 시스템 콜 패턴으로 업데이트하는 단계;를 더 포함하는 악성코드 검출을 위한 자동 규칙 생성방법.

청구항 3

제2항에 있어서,

상기 유사도를 기반으로 그룹화하는 단계는, 상기 업데이트된 악성코드의 시스템 콜 패턴과 기존의 악성코드의 시스템 콜 패턴의 유사성을 비교하여, 설정된 유사성 이상인 악성코드를 하나의 그룹으로 그룹화하는 것을 특징으로 하는 자동 규칙 생성방법.

청구항 4

제1항에 있어서,

상기 각 그룹의 대표 패턴을 선정하여 악성코드 검출 규칙으로 생성하는 단계와 상기 악성코드를 판별하는 단계 사이에,

상기 생성된 악성코드 검출 규칙을 데이터베이스에 저장하는 단계;를 더 포함하는 악성코드 검출을 위한 자동 규칙 생성방법.

청구항 5

제4항에 있어서,

상기 데이터베이스에 저장하는 단계는, 상기 각 그룹의 명칭, 상기 각 그룹의 대표 패턴, 상기 각 그룹에 포함되는 악성코드의 시스템 콜 패턴과 검출 횟수를 매칭시켜 저장하는 것을 특징으로 하는 악성코드 검출을 위한 자동 규칙 생성방법.

청구항 6

제1항에 있어서,

상기 정상 프로그램은, 윈도우즈 OS(operating system)가 설치될 때 디폴트로 설치되는 프로그램 또는 상기 윈도우즈 OS 관리자에 의해 승인된 프로그램을 나타내는 것을 특징으로 하는 악성코드 검출을 위한 자동 규칙 생성방법

청구항 7

제1항에 있어서,

상기 유사도를 기반으로 그룹화하는 단계는, 계층적 클러스터링 알고리즘, 최대 거리 알고리즘, K 평균 알고리즘, Isodata 알고리즘, 또는 로빈-먼로법 알고리즘 중 적어도 하나의 클러스터링 알고리즘을 적용하여 그룹화하는 것을 특징으로 악성코드 검출을 위한 자동 규칙 생성방법.

청구항 8

제1항 내지 제7항 중 어느 한 항의 악성코드 검출을 위한 자동 규칙 생성방법의 각 단계를 컴퓨터 상에서 수행하기 위한 프로그램을 기록한 컴퓨터로 판독 가능한 기록매체.

청구항 9

다수의 악성코드와 정상 프로그램을 실행시켜 다수의 시스템 콜을 추출하는 로그 추출부;

상기 로그 추출부에서 추출된 다수의 시스템 콜 각각이 상기 다수의 악성코드와 상기 정상 프로그램에서 실행된 횟수를 이용하여 점수를 부여하고, 상기 점수 차이가 임계값 이상인 시스템 콜을 추출하고, 상기 다수의 악성코드 중 각각의 악성코드의 추출된 시스템 콜의 패턴을 상기 각각의 악성코드의 시스템 콜 패턴으로 설정하는 패턴 설정부;

상기 각각의 악성코드의 시스템 콜 패턴을 유사도를 기반으로 그룹화하는 그룹 생성부;

상기 그룹화된 악성코드의 시스템 콜 패턴을 다중서열정렬(Multiple sequence alignment) 알고리즘을 이용하여 각 그룹의 대표 패턴을 선정하여 악성코드 검출 규칙으로 생성하는 악성코드 검출 규칙 생성부; 및

상기 생성된 악성코드 검출 규칙과 새로 실행되는 프로그램의 시스템 콜 패턴을 비교하여 악성코드를 판별하는 악성코드 판별부;

를 포함하는 악성코드 검출을 위한 자동 규칙 생성장치.

청구항 10

제9항에 있어서,

상기 패턴 설정부는, 상기 악성코드 판별부에서 상기 새로 실행되는 프로그램을 악성코드인 것으로 판별한 경우, 상기 새로 실행되는 프로그램의 시스템 콜 패턴을 상기 악성코드의 시스템 콜 패턴으로 업데이트하여 설정하는 것을 특징으로 하는 악성코드 검출을 위한 자동 규칙 생성장치.

청구항 11

제10항에 있어서,

상기 그룹 생성부는, 상기 업데이트된 악성코드의 시스템 콜 패턴과 기존의 악성코드의 시스템 콜 패턴의 유사성을 비교하여, 설정된 유사성 이상인 악성코드를 하나의 그룹으로 그룹화하는 것을 특징으로 하는 자동 규칙 생성장치.

청구항 12

제9항에 있어서,

상기 생성된 악성코드 검출 규칙을 저장하는 데이터베이스;를 더 포함하는 악성코드 검출을 위한 자동 규칙 생성장치.

청구항 13

제12항에 있어서,

상기 데이터베이스는, 상기 각 그룹의 명칭, 상기 각 그룹의 대표 패턴, 상기 각 그룹에 포함되는 악성코드의 시스템 콜 패턴과 검출 횟수를 매칭시켜 저장하는 것을 특징으로 하는 악성코드 검출을 위한 자동 규칙 생성장치.

청구항 14

제9항에 있어서,

상기 정상 프로그램은, 윈도우즈 OS(operating system)가 설치될 때 디폴트로 설치되는 프로그램 또는 상기 윈도우즈 OS 관리자에 의해 승인된 프로그램을 나타내는 것을 특징으로 하는 악성코드 검출을 위한 자동 규칙 생성장치.

청구항 15

제9항에 있어서,

상기 그룹 생성부는, 계층적 클러스터링 알고리즘, 최대 거리 알고리즘, K 평균 알고리즘, Isodata 알고리즘, 또는 로빈-먼로법 알고리즘 중 적어도 하나의 클러스터링 알고리즘을 적용하여 그룹화하는 것을 특징으로 악성코드 검출을 위한 자동 규칙 생성장치.

청구항 16

악성코드 판별 어플리케이션을 저장하는 서버; 및

상기 서버로부터 상기 악성코드 판별 어플리케이션을 다운로드 받아 설치하면, 다수의 악성코드와 정상 프로그램을 실행시켜 다수의 시스템 콜을 추출하고, 상기 추출된 다수의 시스템 콜 각각이 상기 다수의 악성코드와 상기 정상 프로그램에서 실행된 횟수를 이용하여 상기 다수의 악성코드 중 각각의 악성코드의 시스템 콜 패턴을 추출하고, 상기 각각의 악성코드의 시스템 콜 패턴을 유사도를 기반으로 그룹화한 후 다중서열정렬(Multiple sequence alignment) 알고리즘을 이용하여 각 그룹의 대표 패턴을 선정하여 악성코드 검출 규칙으로 생성하고, 상기 생성된 악성코드 검출 규칙과 새로 실행되는 프로그램의 시스템 콜 패턴을 비교하여 악성코드를 판별하는 사용자 단말기;

를 포함하는 악성코드 검출을 위한 자동 규칙 생성시스템.

청구항 17

제16항에 있어서,

상기 사용자 단말기는,

상기 새로 실행되는 프로그램이 악성코드인 것으로 판별된 경우, 상기 새로 실행되는 프로그램의 시스템 콜 패턴을 상기 악성코드의 시스템 콜 패턴으로 업데이트한 후, 상기 각 그룹의 대표 패턴을 선정하는 것을 특징으로 하는 악성코드 검출을 위한 자동 규칙 생성시스템.

발명의 설명

기 술 분 야

[0001] 본 발명은 변경되는 악성코드를 검출하기 위한 규칙을 자동 생성하는 악성코드 검출을 위한 자동 규칙 생성방법, 자동 규칙 생성장치, 자동 규칙 생성시스템 및 이를 기록한 컴퓨터로 판독가능한 기록매체에 관한 것이다.

배 경 기 술

[0002] 악성코드를 탐지하기 위해 많이 이용되는 안티바이러스(anti-virus) 소프트웨어는 기본적으로 시그니처(signature) 기반 탐지 기법을 사용한다. 시그니처 기반 탐지 기법은 악성코드 프로그램의 해쉬(hash) 값을 시그니처로 저장한 뒤 의심스러운 프로그램이 실행되면 해쉬 값을 계산하고 저장된 시그니처와 비교하여 악성코드 유무를 판단한다. 그러나, 안티바이러스의 악성코드 탐지 기법에 대해 미리 알고 있는 공격자들은 노풀된 악성코드에 대해 코드 수정을 통해 해쉬 값을 변경하여 바이러스를 유포시킨다. 이러한 경우, 악성코드지만 시그니처로 사용되는 해쉬 값이 변경되엇기 때문에 기존에 저장된 시그니처로 탐지할 수 없게 된다.

[0003] 이러한 단점을 극복하기 위하여, 프로그램에 의해 발생되는 시스템 콜을 추적하고, 추적된 시스템 콜을 분석하여 악성행위를 판별하는 행위 기반 탐지 기법이 사용되고 있다. 행위 기반 탐지 기법의 경우 프로그램에 의해 발생되는 시스템 콜이 무수히 많기 때문에 모든 시스템 콜을 추적하는 방법을 실제 시스템에 적용한다면 시스템 과부하를 발생시킬 수 있다. 또한, 어떤 시스템 콜을 추적할 것인지에 대한 연구가 필요하다. 그러나, 이러한 시스템 콜을 추적하는 방법은 악성 행위로 분류되지 않은 시스템 콜을 중간 중간 호출할 경우, 시그니처가 변경되는 것과 마찬가지로 전체 시스템 콜에 대한 악성 행위에 사용되는 시스템 콜의 비중이 낮아져 악성코드를 검출하기 어려운 문제점이 있다. 따라서, 이러한 변경된 형태의 시스템 콜을 삽입하는 경우 악성코드 검출 방법에 대한 연구가 필요하다.

발명의 내용

해결하려는 과제

[0004] 본 발명은 전술한 문제 및 다른 문제를 해결하기 위하여, 시스템 콜의 패턴을 이용하여 악성코드를 검출하고, 새로운 악성코드의 패턴을 반영하여 악성코드 검출 규칙을 업데이트하는 것을 목적으로 한다.

과제의 해결 수단

[0005] 상기 또는 다른 목적을 달성하기 위해 본 발명의 일 측면에 따르면, 다수의 악성코드와 정상 프로그램을 실행시켜 다수의 시스템 콜을 추출하는 단계, 상기 추출된 다수의 시스템 콜 각각이 상기 다수의 악성코드와 상기 정상 프로그램에서 실행된 횟수를 이용하여 점수를 부여하는 단계, 상기 부여된 점수 차이가 임계값 이상인 시스템 콜을 추출하고, 상기 다수의 악성코드 중 각각의 악성코드의 추출된 시스템 콜의 패턴을 상기 각각의 악성코드의 시스템 콜 패턴으로 설정하는 단계, 상기 각각의 악성코드의 시스템 콜 패턴을 유사도를 기반으로 그룹화하는 단계, 상기 그룹화된 악성코드의 시스템 콜 패턴을 다중서열정렬(Multiple sequence alignment) 알고리즘을 이용하여 각 그룹의 대표 패턴을 선정하여 악성코드 검출 규칙으로 생성하는 단계, 및 상기 생성된 악성코드 검출 규칙과 새로 실행되는 프로그램의 시스템 콜 패턴을 비교하여 악성코드를 판별하는 단계를 포함하는 악성

코드 검출을 위한 자동 규칙 생성방법을 제공한다.

[0006] 또한, 본 발명에 따른 악성코드 검출을 위한 자동 규칙 생성방법은 악성코드를 판별하는 단계 다음에, 상기 새로 실행되는 프로그램이 악성코드인 것으로 판별된 경우, 상기 새로 실행되는 프로그램의 시스템 콜 패턴을 상기 악성코드의 시스템 콜 패턴으로 업데이트하는 단계를 더 포함할 수 있다.

[0007] 유사도를 기반으로 그룹화하는 단계는, 상기 업데이트된 악성코드의 시스템 콜 패턴과 기존의 악성코드의 시스템 콜 패턴의 유사성을 비교하여, 설정된 유사성 이상인 악성코드를 하나의 그룹으로 그룹화할 수 있다.

[0008] 악성코드 검출을 위한 자동 규칙 생성방법은 각 그룹의 대표 패턴을 선정하여 악성코드 검출 규칙으로 생성하는 단계와 상기 악성코드를 판별하는 단계 사이에, 상기 생성된 악성코드 검출 규칙을 데이터베이스에 저장하는 단계를 더 포함할 수 있다.

[0009] 데이터베이스에 저장하는 단계는, 상기 각 그룹의 명칭, 상기 각 그룹의 대표 패턴, 상기 각 그룹에 포함되는 악성코드의 시스템 콜 패턴과 검출 횟수를 매칭시켜 저장할 수 있다.

[0010] 정상 프로그램은 윈도우즈 OS(operating system)가 설치될 때 디폴트로 설치되는 프로그램 또는 상기 윈도우즈 OS 관리자에 의해 승인된 프로그램을 나타낼 수 있다.

[0011] 유사도를 기반으로 그룹화하는 단계는 계층적 클러스터링 알고리즘, 최대 거리 알고리즘, K 평균 알고리즘, Isodata 알고리즘, 또는 로빈-먼로법 알고리즘 중 적어도 하나의 클러스터링 알고리즘을 적용하여 그룹화할 수 있다.

[0012] 또한, 상기 또는 다른 목적을 달성하기 위해 본 발명의 다른 측면에 따르면, 상기악성코드 검출을 위한 자동 규칙 생성방법의 각 단계를 컴퓨터 상에서 수행하기 위한 프로그램을 기록한 컴퓨터로 판독 가능한 기록매체를 제공한다.

[0013] 또한, 상기 또는 다른 목적을 달성하기 위해 본 발명의 다른 측면에 따르면, 다수의 악성코드와 정상 프로그램을 실행시켜 다수의 시스템 콜을 추출하는 로그 추출부, 상기 로그 추출부에서 추출된 다수의 시스템 콜 각각이 상기 다수의 악성코드와 상기 정상 프로그램에서 실행된 횟수를 이용하여 점수를 부여하고, 상기 점수 차이가 임계값 이상인 시스템 콜을 추출하고, 상기 다수의 악성코드 중 각각의 악성코드의 추출된 시스템 콜의 패턴을 상기 각각의 악성코드의 시스템 콜 패턴으로 설정하는 패턴 설정부, 상기 각각의 악성코드의 시스템 콜 패턴을 유사도를 기반으로 그룹화하는 그룹 생성부, 상기 그룹화된 악성코드의 시스템 콜 패턴을 다중서열정렬(Multiple sequence alignment) 알고리즘을 이용하여 각 그룹의 대표 패턴을 선정하여 악성코드 검출 규칙으로 생성하는 악성코드 검출 규칙 생성부, 및 상기 생성된 악성코드 검출 규칙과 새로 실행되는 프로그램의 시스템 콜 패턴을 비교하여 악성코드를 판별하는 악성코드 판별부를 포함하는 악성코드 검출을 위한 자동 규칙 생성장치를 제공한다.

[0014] 패턴 설정부는 상기 악성코드 판별부에서 상기 새로 실행되는 프로그램을 악성코드인 것으로 판별한 경우, 상기 새로 실행되는 프로그램의 시스템 콜 패턴을 상기 악성코드의 시스템 콜 패턴으로 업데이트하여 설정할 수 있다.

[0015] 그룹 생성부는 상기 업데이트된 악성코드의 시스템 콜 패턴과 기존의 악성코드의 시스템 콜 패턴의 유사성을 비교하여, 설정된 유사성 이상인 악성코드를 하나의 그룹으로 그룹화할 수 있다.

[0016] 또한, 상기 또는 다른 목적을 달성하기 위해 본 발명의 다른 측면에 따르면, 악성코드 판별 어플리케이션을 저장하는 서버, 및 상기 서버로부터 상기 악성코드 판별 어플리케이션을 다운로드 받아 설치하면, 다수의 악성코드와 정상 프로그램을 실행시켜 다수의 시스템 콜을 추출하고, 상기 추출된 다수의 시스템 콜 각각이 상기 다수의 악성코드와 상기 정상 프로그램에서 실행된 횟수를 이용하여 상기 다수의 악성코드 중 각각의 악성코드의 시스템 콜 패턴을 추출하고, 상기 각각의 악성코드의 시스템 콜 패턴을 유사도를 기반으로 그룹화한 후 다중서열정렬(Multiple sequence alignment) 알고리즘을 이용하여 각 그룹의 대표 패턴을 선정하여 악성코드 검출 규칙으로 생성하고, 상기 생성된 악성코드 검출 규칙과 새로 실행되는 프로그램의 시스템 콜 패턴을 비교하여 악성코드를 판별하는 사용자 단말기를 포함하는 악성코드 검출을 위한 자동 규칙 생성시스템을 제공한다.

발명의 효과

[0017] 본 발명에 따른 악성코드 검출을 위한 자동 규칙 생성방법, 자동 규칙 생성장치, 자동 규칙 생성시스템 및 이를 기록한 컴퓨터로 판독가능한 기록매체의 효과에 대해 설명하면 다음과 같다.

[0018] 본 발명의 실시 예들 중 적어도 하나에 의하면, 시스템 콜의 패턴을 이용하여 악성코드를 검출함으로써 악성코드에 사용되지 않던 시스템 콜이 중간에 삽입되는 경우에도 패턴을 이용하여 악성코드를 검출할 수 있다는 장점이 있다.

[0019] 또한, 본 발명의 실시 예들 중 적어도 하나에 의하면, 새로운 악성코드의 시스템 콜 패턴을 이용하여 악성코드의 시스템 콜의 패턴을 업데이트함으로써 악성코드 검출에 적응적으로 대응할 수 있다는 장점이 있다.

[0020] 본 발명의 적용 가능성의 추가적인 범위는 이하의 상세한 설명으로부터 명백해질 것이다. 그러나 본 발명의 사상 및 범위 내에서 다양한 변경 및 수정은 당업자에게 명확하게 이해될 수 있으므로, 상세한 설명 및 본 발명의 바람직한 실시 예와 같은 특정 실시 예는 단지 예시로 주어진 것으로 이해되어야 한다.

도면의 간단한 설명

[0021] 도 1 내지 도 4는 본 발명의 일 실시예와 관련된 악성코드 검출을 위한 자동 규칙 생성방법의 흐름도들이다.

도 5는 본 발명의 일 실시예와 관련된 악성코드 검출을 위한 자동 규칙 생성장치의 개략적인 구성도이다.

도 6은 본 발명의 일 실시예와 관련된 악성코드 검출을 위한 자동 규칙 생성시스템의 개략적인 구성도이다.

도 7은 본 발명의 일 실시예와 관련된 악성코드 검출 방법을 설명하기 위한 예시도이다.

발명을 실시하기 위한 구체적인 내용

[0022] 이하, 첨부된 도면을 참조하여 본 명세서에 개시된 실시 예를 상세히 설명하되, 도면 부호에 관계없이 동일하거나 유사한 구성요소는 동일한 참조 번호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다. 이하의 설명에서 사용되는 구성요소에 대한 접미사 "모듈" 및 "부"는 명세서 작성의 용이함만이 고려되어 부여되거나 혼용되는 것으로서, 그 자체로 서로 구별되는 의미 또는 역할을 갖는 것은 아니다. 또한, 본 명세서에 개시된 실시 예를 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 명세서에 개시된 실시 예의 요지를 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다. 또한, 첨부된 도면은 본 명세서에 개시된 실시 예를 쉽게 이해할 수 있도록 하기 위한 것일 뿐, 첨부된 도면에 의해 본 명세서에 개시된 기술적 사상이 제한되지 않으며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0023] 제1, 제2 등과 같이 서수를 포함하는 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되지는 않는다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.

[0024] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다.

[0025] 본 출원에서, "포함한다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0026] 본 명세서에서 설명되는 악성코드 검출을 위한 자동 규칙 생성장치 또는 사용자 단말기는 태블릿 PC(tablet PC), 울트라북(ultrabook), 랩탑, 컴퓨터, 휴대폰, 스마트 폰(smart phone), 디지털방송용 단말기, PDA(personal digital assistants), PMP(portable multimedia player) 등의 OS(Operating System)이 동작가능한 전자 장치를 포함할 수 있다.

[0028] 도 1 내지 도 4는 본 발명의 일 실시예와 관련된 악성코드 검출을 위한 자동 규칙 생성방법의 흐름도들이다.

[0029] 도 1을 참조하면, 본 발명의 일 실시예와 관련된 악성코드 검출을 위한 자동 규칙 생성방법은 정상 프로그램과 악성코드의 시스템 콜을 추출하고(S110), 악성코드의 시스템 콜 패턴을 설정하고(S120), 악성코드들을 그룹화할 수 있다(S130).

[0030] 구체적으로, 정상 프로그램과 악성코드의 시스템 콜을 추출하는 단계(S110)는 다수의 악성코드와 정상 프로그램

을 샘플로 추출하여 실행시켜 다수의 시스템 콜을 추출할 수 있다. 여기서, 악성 프로그램은 특정 파일을 암호화한 후 삭제하거나 서로 다른 파일을 암호화된 하나의 파일로 변경하는 등 윈도우즈 OS의 정상적인 작동을 방해하는 프로그램을 나타내고, 정상 프로그램은 윈도우즈 OS(operating system)가 설치될 때 디폴트로 설치되는 프로그램 또는 윈도우즈 OS 관리자에 의해 승인된 프로그램을 나타낸다. 예를 들어, 악성 프로그램은 랜섬웨어, 트로이 목마 등이 될 수 있고, 정상 프로그램은 윈도우즈 OS에 기본적으로 설치된 메모장, 윈도우즈 익스플로러 등이 될 수 있다.

[0031] 다음으로, 악성코드의 시스템 콜 패턴을 설정하는 단계(S120)는 정상 프로그램의 시스템 콜과 악성코드의 시스템 콜의 수행 횟수를 비교하여 정상 코드의 시스템 콜 패턴인지 악성코드의 시스템 콜 패턴인지 설정할 수 있다.

[0032] 도 2를 참조하면, 정상 프로그램 또는 악성코드가 실행되면 시스템 콜의 실행 횟수를 확인하고(S122), 각 시스템 콜에 점수를 부여할 수 있다(S124). 구체적으로, "A"라는 시스템 콜이 악성코드에서 10회 사용되고, 정상 프로그램에서 2회 사용된 경우, "A" 시스템 콜에 8이라는 점수를 부여하고, "B"라는 시스템 콜이 악성코드에서 4회 사용되고, 정상 프로그램에서 15회 사용된 경우, "B" 시스템 콜에 -11이라는 점수를 부여하는 방법으로 각각의 시스템 콜에 점수를 부여할 수 있다.

[0033] 부여된 점수가 임계값 이상인지 판단하고(S126), 임계값 이상인 경우, 각각의 악성 코드의 시스템 콜 패턴으로 설정하고(S128), 임계값 미만인 경우, 정상 코드의 시스템 콜 패턴으로 설정할 수 있다(S128'). 이때, 임계값은 기준값을 나타내며, 기준값이 음수 (-)인 경우 절대값이 큰 경우 임계값 이상으로 판단할 수 있다. 구체적으로, "A" 시스템 콜의 점수가 8이고 "B" 시스템 콜의 점수가 -11인 경우, "A" 시스템 콜은 "B" 시스템 콜에 비하여 악성코드에서 사용될 가능성이 더 높고, 각각의 시스템 콜의 점수를 이용하여 악성코드에 사용될 가능성이 더 높은 시스템 콜을 추출할 수 있다. 즉, 각각의 시스템 콜의 점수와 기준 점수를 비교하여, 악성코드에 사용될 가능성이 더 높은 시스템 콜을 검출할 수 있다.

[0034] 위와 같은 방법으로, 악성코드에 임계값 이상인 시스템 콜을 순차적으로 검출하고, 검출된 시스템 콜을 나열하여 사용 패턴을 검출할 수 있다. 검출된 사용 패턴을 악성코드의 시스템 콜 패턴으로 설정할 수 있다.

[0036] 다음으로, 각각의 악성코드의 시스템 콜 패턴을 그룹화하는 단계(S130)는 각 시스템 콜 패턴의 유사도를 기반으로 그룹화를 수행할 수 있다. 구체적으로, 계층적 클러스터링 알고리즘, 최대 거리 알고리즘, K 평균 알고리즘, Isodata 알고리즘, 또는 로빈-먼로법 알고리즘 중 적어도 하나의 클러스터링 알고리즘을 적용하여 그룹화를 수행할 수 있다.

[0037] 설정된 악성코드 시스템 콜 중 공통으로 포함된 악성코드 시스템 콜의 개수에 따라 유사도가 계산될 수 있다. 계산된 유사도에 따라 악성코드들을 그룹화할 수 있다.

[0038] 다음으로, 본 발명에 따른 악성코드 검출을 위한 자동 규칙 생성방법은 악성코드 검출 규칙을 생성하고(S140), 생성된 규칙에 따라 악성코드를 판별할 수 있다(S150).

[0039] 구체적으로, 악성코드 검출 규칙을 생성하는 단계(S140)는 각각의 그룹 별로 대표 악성코드 시스템 콜의 패턴을 결정할 수 있다. 이때, 그룹화된 악성코드의 시스템 콜 패턴을 다중서열정렬(Multiple sequence alignment) 알고리즘을 이용하여 각 그룹의 대표 패턴을 선정하여 각 그룹의 악성코드 검출 규칙으로 결정할 수 있다. 다중서열정렬 알고리즘은 악성코드의 시스템 콜 패턴이 여러 개 있을때 한꺼번에 묶에서 서로 정렬하는 방법을 의미한다.

[0040] 생성된 악성코드 검출 규칙(즉, 각 그룹의 대표 패턴)과 새로 실행되는 프로그램의 시스템 콜 패턴을 비교하여 악성코드인지 여부와 어느 그룹에 포함되는지 결정할 수 있다. 예를 들어, 제1 그룹의 악성코드의 시스템 콜 패턴이 A->D->C->D로 설정된 경우, 특정 프로그램이 A->F->D->a->b->C->D로 시스템 콜이 수행되면 설정된 악성코드의 시스템 콜 패턴인 A->D->C->D 순서로 시스템 콜이 수행되고 있으므로, 악성코드로 판단할 수 있다. 즉, 특정 프로그램은 제1 그룹의 악성코드로 판단될 수 있다. 또한, 유사도에 따라 임계값 이상의 유사도를 만족하는 경우, 제1 그룹의 악성코드로 판단할 수 있다.

[0041] 도 7을 참조하면, 각각의 악성코드의 시스템 콜에서, 악성코드에 사용될 가능성이 높은 시스템 콜 패턴이 A->D->E->E->D->B으로 설정된 경우, 각 악성코드 시스템 콜의 시퀀스에 따라 유사도를 계산할 수 있다.

[0042] 제1 악성코드는 시퀀스 I의 순서로 즉, C, A, D, C, E, E, D, B, C, C 순서의 시스템 콜 10개로 실행되고, 제2 악성코드는 A, D, E, E, D, B, D, C, A, C 순서의 시스템 콜 10개로 실행될 수 있다.

[0043] 이 경우, 제1 악성코드와 제2 악성코드는 설정된 패턴과 각각 0.6의 유사도를 갖는것으로 계산될 수 있고, 미리 설정된 유사도가 0.5인 경우, 악성코드로 판별될 수 있다.

[0045] 도 3을 참조하면, 본 발명의 일 실시예에 따른 악성코드 검출을 위한 자동 규칙 생성방법은 새로 실행되는 프로그램이 악성코드인 것으로 판별된 경우(S150), 새로 실행되는 프로그램의 시스템 콜 패턴을 악성코드의 시스템 콜 패턴으로 업데이트할 수 있다(S160).

[0046] 구체적으로, 악성코드 유포자가 악성코드의 시스템 콜을 A->D->C->B 패턴에서 A->D->b->a->C->B 패턴으로 변경하여 유포한 경우, 변경된 악성코드 역시 악성코드의 시스템 콜 패턴으로 설정된 A->D->C->B 패턴을 포함하고 있기 때문에 악성코드로 검출되지만, 계속적으로 변형된 악성코드가 유포될 경우, 일부 시스템 콜을 변경하게 되면 악성코드로 검출되지 않을 수 있다. 따라서, 변경된 악성코드의 시스템 콜 패턴인 A->D->b->a->C->B 을 악성코드의 시스템 콜 패턴으로 업데이트 할 수 있다.

[0047] 다음으로, 업데이트된 악성코드의 시스템 콜 패턴과 기존의 악성코드의 시스템 콜 패턴의 유사성을 비교하여, 설정된 유사성 이상인 악성코드를 하나의 그룹으로 그룹화할 수 있다(S170). 구체적으로, 변경된 악성코드의 시스템 콜 패턴인 A->D->b->a->C->B 을 각 그룹의 대표 패턴들과 비교하여 특정 그룹으로 분류하고, 특정 그룹에 포함된 다수의 악성코드의 시스템 콜 패턴을 다중서열정렬 알고리즘을 다시 적용하여 대표 패턴을 결정할 수 있다.

[0048] 위와 같은 방법에 의하여, 변형된 악성코드의 시스템 콜 패턴이 악성코드의 대표 패턴을 설정하는데 연속적으로 반영될 수 있다.

[0050] 도 4를 참조하면, 본 발명의 일 실시예에 따른 악성코드 검출을 위한 자동 규칙 생성방법은 각 그룹의 대표 패턴을 선정하여 악성코드 검출 규칙으로 생성하는 단계(S140)와 악성코드를 판별하는 단계(S150) 사이에, 생성된 악성코드 검출 규칙을 데이터베이스에 저장하는 단계(S145)를 더 포함할 수 있다.

[0051] 구체적으로, 악성코드를 분류한 각 그룹의 명칭, 각 그룹의 대표 패턴, 각 그룹에 포함되는 악성코드의 시스템 콜 패턴과 검출 횟수를 매칭시켜 저장할 수 있고, 새로운 악성코드가 검출된 경우, 새로운 악성코드의 변경된 시스템 콜 패턴과 분류된 그룹 등을 데이터베이스에 업데이트 할 수 있다.

[0053] 도 5는 본 발명의 일 실시예와 관련된 악성코드 검출을 위한 자동 규칙 생성장치의 개략적인 구성도이다.

[0054] 도 5를 참조하면, 본 발명의 일 실시예와 관련된 악성코드 검출을 위한 자동 규칙 생성장치(100)는 로그 추출부(110), 패턴 설정부(120), 그룹 생성부(130), 악성코드 검출 규칙 생성부(140), 및 악성코드 판별부(150)를 포함하여 구성될 수 있다. 또한, 악성코드 검출을 위한 자동 규칙 생성장치(100)는 악성코드와 관련된 정보를 저장하는 데이터베이스(DB)(160)를 더 포함하여 구성될 수 있다.

[0055] 로그 추출부(110)는 다수의 악성코드와 정상 프로그램을 실행시켜 다수의 시스템 콜을 추출할 수 있다. 즉, 샘플로 선정된 각각의 악성코드와 정상 프로그램을 실행시켜 시스템 콜을 추출하고, 각 시스템 콜의 수행 횟수와 수행 순서를 추출할 수 있다.

[0056] 패턴 생성부(120)는 로그 추출부(110)에서 추출된 다수의 시스템 콜 각각이 다수의 악성코드와 정상 프로그램에서 실행된 횟수를 이용하여 점수를 부여할 수 있다. 또한, 패턴 생성부(120)는 점수 차이가 임계값 이상인 시스템 콜을 추출하고, 다수의 악성코드 중 각각의 악성코드의 추출된 시스템 콜의 패턴을 각각의 악성코드의 시스템 콜 패턴으로 설정할 수 있다. 시스템 콜 패턴을 설정하는 구체적인 방법은 도 2에서 상세히 설명하였다.

[0057] 그룹 생성부(130)는 각각의 악성코드의 시스템 콜 패턴을 유사도를 기반으로 그룹화할 수 있다. 각 그룹에 포함된 악성코드의 시스템 콜 패턴은 설정된 기준값 이상의 유사도를 갖도록 분류될 수 있다. 그룹 생성부(130)는 계층적 클러스터링 알고리즘, 최대 거리 알고리즘, K 평균 알고리즘, Isodata 알고리즘, 또는 로빈-먼로법 알고리즘 중 적어도 하나의 클러스터링 알고리즘을 적용하여 그룹화할 수 있다.

[0058] 악성코드 검출 규칙 생성부(140)는 그룹화된 악성코드의 시스템 콜 패턴을 다중서열정렬(Multiple sequence alignment) 알고리즘을 이용하여 각 그룹의 대표 패턴을 선정하여 악성코드 검출 규칙으로 생성할 수 있다. 하나의 그룹에 포함된 다수의 악성코드들의 시스템 콜 패턴을 나열하고, 다수의 시스템 콜 패턴을 이용하여 각 그룹의 악성코드들의 대표 시스템 콜 패턴을 결정하여 악성코드 검출 규칙으로 설정할 수 있다.

[0059] 악성코드 판별부(150)는 생성된 악성코드 검출 규칙과 새로 실행되는 프로그램의 시스템 콜 패턴을 비교하여 악성코드를 판별할 수 있다. 이때, 악성코드 판별부(150)는 미리 설정된 유사도 이상 일치하는 경우 악성코드로

판별할 수 있다. 이 경우, 악성코드 판별부(150)는 각 그룹의 대표 패턴과 유사도를 비교하여 악성코드로 판별된 경우, 해당 그룹의 악성코드로 재분류할 수 있다.

[0060] 또한, 패턴 설정부(120)는 악성코드 판별부(150)에서 새로 실행되는 프로그램을 악성코드인 것으로 판별한 경우, 새로 실행되는 프로그램의 시스템 콜 패턴을 악성코드의 시스템 콜 패턴으로 업데이트하여 설정할 수 있다. 즉, 해당 그룹의 악성코드 패턴 중 하나로 포함시키도록 데이터베이스를 업데이트하고, 새로운 악성코드 패턴이 포함된 풀에서 대표 패턴을 결정할 수 있다. 또한, 검출된 악성코드를 해당하는 그룹에 포함시키지 않고, 그룹 생성부(130)에서 업데이트된 악성코드의 시스템 콜 패턴과 기존의 악성코드의 시스템 콜 패턴의 유사성을 비교하여, 설정된 유사성 이상인 악성코드를 하나의 그룹으로 그룹화한 후 풀이 변경된 그룹의 대표 패턴을 새로 결정할 수 있다.

[0061] 데이터베이스(160)는 생성된 악성코드 검출 규칙을 저장할 수 있다. 이때, 데이터베이스(160)는 각 그룹의 명칭, 각 그룹의 대표 패턴, 각 그룹에 포함되는 악성코드의 시스템 콜 패턴과 검출 횟수를 매칭시켜 저장할 수 있다.

[0063] 도 6은 본 발명의 일 실시예와 관련된 악성코드 검출을 위한 자동 규칙 생성시스템의 개략적인 구성도이다.

[0064] 도 6을 참조하면, 본 발명의 일 실시예와 관련된 악성코드 검출을 위한 자동 규칙 생성시스템(1000)은 적어도 하나의 사용자 단말기(100A, 100B, 100C) 및 서버(200)를 포함하여 구성될 수 있다.

[0065] 먼저, 서버(200)는 악성코드 판별 어플리케이션을 저장하고, 사용자 단말기(100A, 100B, 100C)와 유무선으로 연결되어, 악성코드 판별 어플리케이션의 다운로드 요청을 수신하는 경우, 악성코드 판별 어플리케이션을 사용자 단말기(100A, 100B, 100C)로 전송할 수 있다.

[0066] 사용자 단말기(100A, 100B, 100C)는 서버(200)로부터 악성코드 판별 어플리케이션을 요청하고, 다운로드 받아 설치할 수 있다. 또한, 사용자 단말기(100A, 100B, 100C)는 악성코드 판별 어플리케이션을 설치하면, 다수의 악성코드와 정상 프로그램을 실행시켜 다수의 시스템 콜을 추출하고, 추출된 다수의 시스템 콜 각각이 다수의 악성코드와 정상 프로그램에서 실행된 횟수를 이용하여 다수의 악성코드 중 각각의 악성코드의 시스템 콜 패턴을 추출할 수 있다. 다음으로, 사용자 단말기(100A, 100B, 100C)는 각각의 악성코드의 시스템 콜 패턴을 유사도를 기반으로 그룹화한 후 다중서열정렬(Multiple sequence alignment) 알고리즘을 이용하여 각 그룹의 대표 패턴을 선정하여 악성코드 검출 규칙으로 생성하고, 생성된 악성코드 검출 규칙과 새로 실행되는 프로그램의 시스템 콜 패턴을 비교하여 악성코드를 판별할 수 있다.

[0067] 또한, 악성코드 검출 어플리케이션이 설치된 사용자 단말기(100A, 100B, 100C)는 새로 실행되는 프로그램이 악성코드인 것으로 판별된 경우, 새로 실행되는 프로그램의 시스템 콜 패턴을 악성코드의 시스템 콜 패턴으로 업데이트한 후, 각 그룹의 대표 패턴을 선정할 수 있다. 즉, 사용자 단말기(100A, 100B, 100C)는 상기 악성코드 검출을 위한 자동 규칙 생성장치로 기능할 수 있다.

[0069] 본 발명에 따르면, 시스템 콜의 패턴을 이용하여 악성코드를 검출함으로써 악성코드에 사용되지 않던 시스템 콜이 중간에 삽입되는 경우에도 패턴을 이용하여 악성코드를 검출할 수 있고, 새로운 악성코드의 시스템 콜 패턴을 이용하여 악성코드의 시스템 콜의 패턴을 업데이트함으로써 악성코드 검출에 적응적으로 대응할 수 있다.

[0071] 상기 또는 다른 목적을 달성하기 위해 본 발명의 다른 측면에 따르면, 악성코드 검출을 위한 자동 규칙 생성방법의 각 단계를 컴퓨터 상에서 수행하기 위한 프로그램을 기록한 컴퓨터로 판독 가능한 기록매체를 포함한다.

[0072] 전술한 본 발명은, 프로그램이 기록된 매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 매체는, 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 매체의 예로는, HDD(Hard Disk Drive), SSD(Solid State Disk), SDD(Silicon Disk Drive), ROM, RAM, CD-ROM, 자기 테이프, 플로피 디스크, 광 데이터 저장 장치 등을 포함한다. 또한, 상기 컴퓨터는 단말기의 제어부(180)를 포함할 수도 있다. 따라서, 상기의 상세한 설명은 모든 면에서 제한적으로 해석되어서는 아니되고 예시적인 것으로 고려되어야 한다. 본 발명의 범위는 첨부된 청구항의 합리적 해석에 의해 결정되어야 하고, 본 발명의 등가적 범위 내에서의 모든 변경은 본 발명의 범위에 포함된다.

부호의 설명

[0073] 1000: 악성코드 검출을 위한 자동 규칙 생성시스템

100A, 100B, 100C: 사용자 단말기 200: 서버

100: 악성코드 검출을 위한 자동 규칙 생성장치

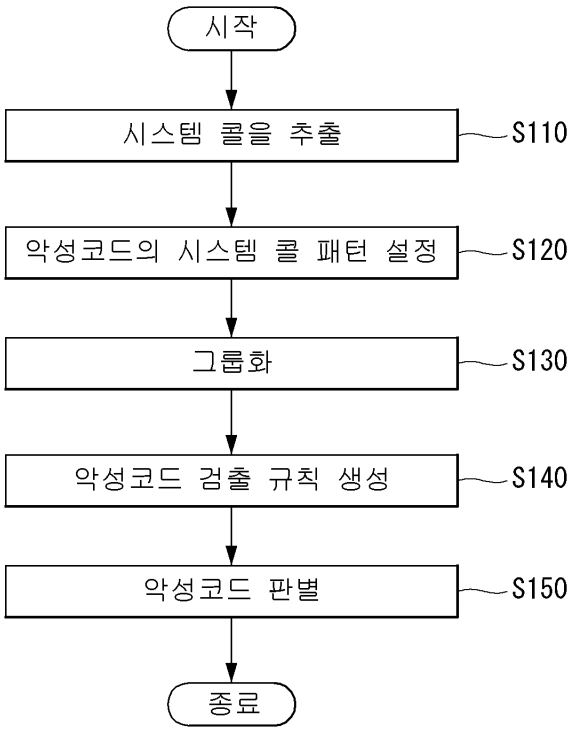
110: 로그 추출부 120: 패턴 설정부

130: 그룹 생성부 140: 악성코드 검출 규칙 생성부

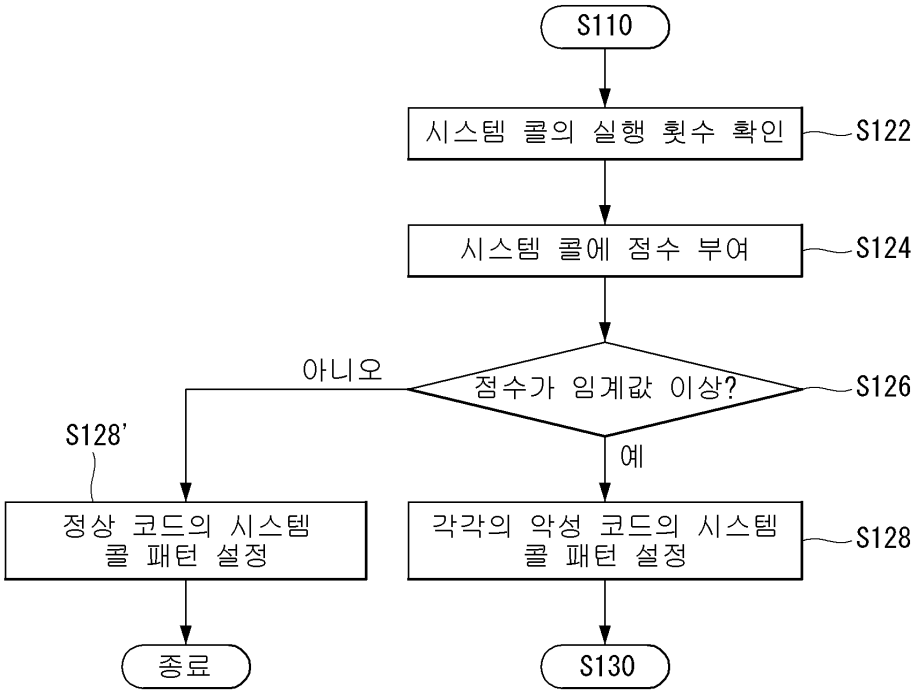
150: 악성코드 판별부 160: 데이터베이스(DB)

도면

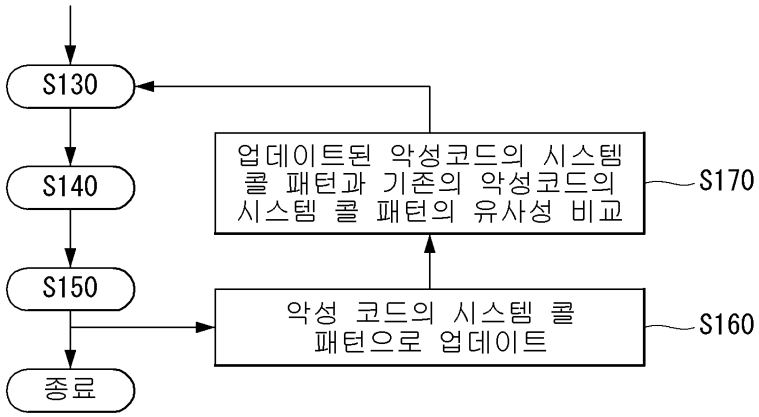
도면1



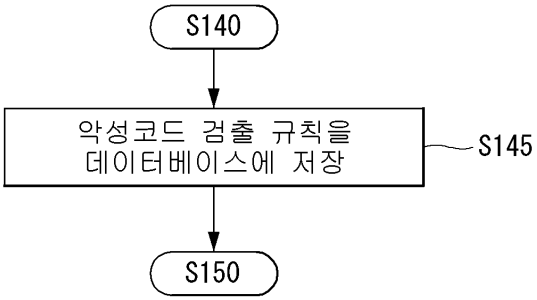
도면2


S110
시스템 콜의 실행 횟수 확인
S122
시스템 콜에 점수 부여
S124
점수가 임계값 이상?
S126
아니오
예
S128'
정상 코드의 시스템 콜 패턴 설정
각각의 악성 코드의 시스템 콜 패턴 설정
S128
종료
S130

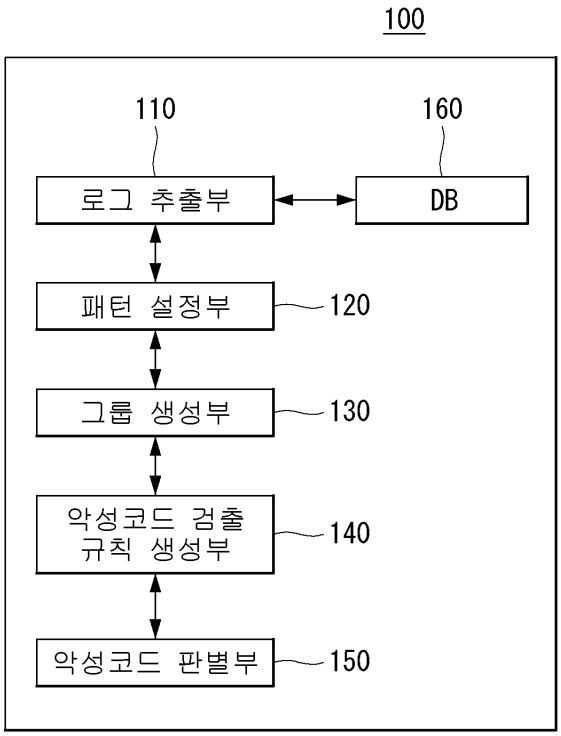
도면3


S130
S140
S150
종료
업데이트된 악성코드의 시스템 콜 패턴과 기존의 악성코드의 시스템 콜 패턴의 유사성 비교
S170
악성 코드의 시스템 콜 패턴으로 업데이트
S160

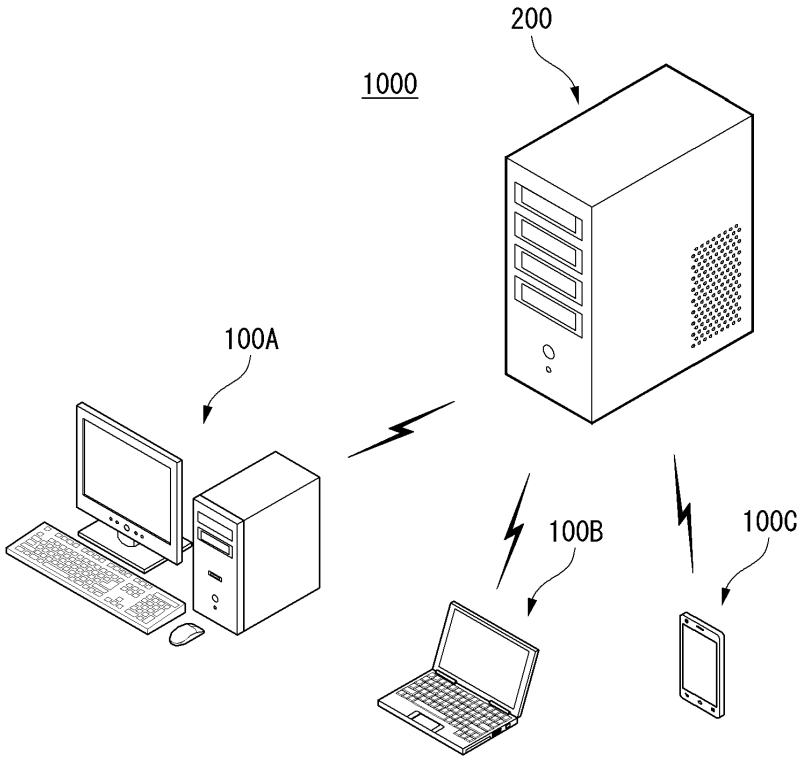
도면4


S140
악성코드 검출 규칙을 데이터베이스에 저장
S145
S150

도면5


100
110
160
로그 추출부
DB
패턴 설정부
120
그룹 생성부
130
악성코드 검출
규칙 생성부
140
악성코드 판별부
150

도면6


200
1000
100A
100B
100C

도면7