(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년02월28일
(11) 등록번호 10-2369240
(24) 등록일자 2022년02월24일

(51) 국제특허분류(Int. Cl.)
H04L 9/40 (2022.01) *G06N 3/04* (2006.01)
G06N 3/08 (2006.01) *H04L 43/00* (2022.01)
H04L 47/00 (2022.01)
(52) CPC특허분류
H04L 63/1425 (2013.01)
G06N 3/0454 (2013.01)
(21) 출원번호 10-2020-0103350
(22) 출원일자 2020년08월18일
심사청구일자 2020년08월18일
(65) 공개번호 10-2022-0022322
(43) 공개일자 2022년02월25일
(56) 선행기술조사문헌
JP2019047335 A*
KR1020200087299 A*
KR102074909 B1*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
(72) 발명자
윤주범
서울특별시 송파구 충민로4길 19, 704동 401호(장지동, 송파파인타운7단지)
이영우
경기도 구리시 경춘로288번길 39, 마동 410호(수택동)
(뒷면에 계속)
(74) 대리인
두호특허법인

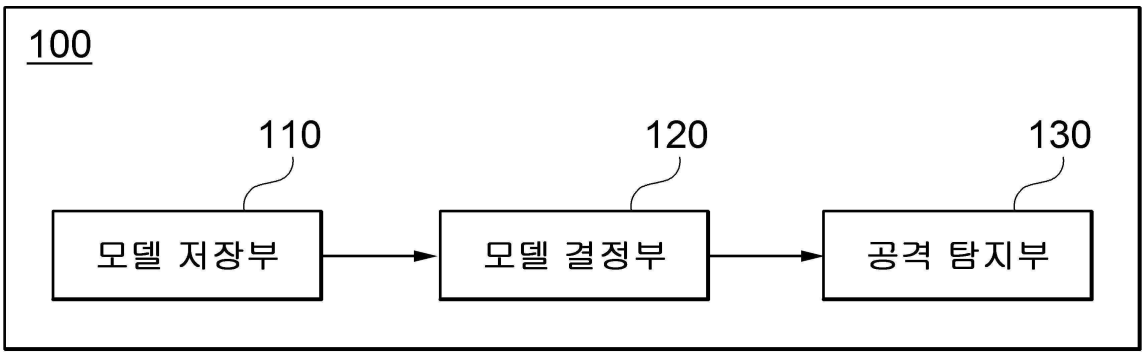
전체 청구항 수 : 총 16 항 심사관 : 문형섭

(54) 발명의 명칭 **네트워크 공격 탐지 장치 및 방법**

(57) 요 약

네트워크 공격 탐지 장치 및 방법이 개시된다. 일 실시예에 따른 네트워크 공격 탐지 장치는, 각각 복수의 정상 네트워크 데이터 셋 중 하나와 복수의 딥 러닝(Deep learning) 알고리즘 중 하나를 이용한 학습을 통해 생성된 복수의 네트워크 특성 분류 모델을 저장하는 모델 저장부, 상기 복수의 네트워크 특성 분류 모델 중 분류 대상인 네트워크 특성 추출 파일에 대한 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정하는 모델 결정부 및 상기 결정된 네트워크 특성 분류 모델을 이용하여 상기 네트워크 특성 추출 파일에 대해 상기 정상-비정상 분류를 수행하고, 상기 정상-비정상 분류의 결과에 기초하여 상기 네트워크 특성 추출 파일에 대응되는 네트워크에 대한 공격을 탐지하는 공격 탐지부를 포함한다.

대 표 도 - 도1



(52) CPC특허분류
G06N 3/08 (2013.01)
H04L 43/026 (2013.01)
H04L 43/0829 (2013.01)
H04L 43/16 (2013.01)
H04L 47/2441 (2013.01)

(72) 발명자
조관용
서울특별시 노원구 마들로 127, 37동 909호(월계동, 월계삼호아파트)
박준영
서울특별시 도봉구 도봉로136길 28, 508동 2304호(창동, 북한산 아이파크)
손배훈
충청남도 서산시 향교4로 8-13, 104동 1101호(동문동, 현진에버빌)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711120086
과제번호	2020-0-01602-001
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	대학ICT연구센터지원사업
연구과제명	지능형 사이버 위협 대응 기술 개발 및 인력양성
기 여 율	1/1
과제수행기관명	숭실대학교 산학협력단
연구기간	2020.07.01 ~ 2021.12.31

명 세 서

청구범위

청구항 1

각각 복수의 정상 네트워크 데이터 셋 중 하나와 복수의 딥 러닝(Deep learning) 알고리즘 중 하나를 이용한 학습을 통해 생성된 복수의 네트워크 특성 분류 모델을 저장하는 모델 저장부;

상기 복수의 네트워크 특성 분류 모델 중 분류 대상인 네트워크 특성 추출 파일에 대한 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정하는 모델 결정부; 및

상기 결정된 네트워크 특성 분류 모델을 이용하여 상기 네트워크 특성 추출 파일에 대해 상기 정상-비정상 분류를 수행하고, 상기 정상-비정상 분류의 결과에 기초하여 상기 네트워크 특성 추출 파일에 대응되는 네트워크에 대한 공격을 탐지하는 공격 탐지부를 포함하되,

상기 모델 결정부는,

정상-비정상 레이블(label)로 라벨링(labeling)된 복수의 테스트 파일을 포함하는 테스트 데이터 셋 및 상기 정상-비정상 분류의 기준이 되는 기 설정된 임계 값에 기초하여 상기 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정하는, 네트워크 공격 탐지 장치.

청구항 2

청구항 1에 있어서,

상기 복수의 딥 러닝 알고리즘은,

오토 인코더(Auto Encoder) 알고리즘 및 순환 신경망(RNN; Recurrent Neural Network) 알고리즘 중 적어도 하나를 포함하는, 네트워크 공격 탐지 장치.

청구항 3

청구항 1에 있어서,

상기 공격 탐지부에 입력되거나 상기 복수의 정상 네트워크 데이터 셋 중 어느 하나에 포함된 복수의 네트워크 패킷 캡쳐 파일(network packet capture file)을 하나 이상의 네트워크 패킷 플로우(network packet flow)로 군집화하고, 기 설정된 종류의 네트워크 특성에 기초하여 상기 하나 이상의 네트워크 패킷 플로우 각각에 대응되는 네트워크 특성 추출 파일을 생성하는 전처리부를 더 포함하는, 네트워크 공격 탐지 장치.

청구항 4

삭제

청구항 5

청구항 1에 있어서,

상기 모델 결정부는,

상기 임계 값에 기초한 상기 복수의 테스트 파일의 정상-비정상 분류 결과와 상기 정상-비정상 레이블에 기초하여 상기 복수의 네트워크 특성 분류 모델 각각의 정밀도(Precision) 및 재현율(Recall)을 산출하고, 상기 정밀도 및 상기 재현율에 기초하여 F1-score를 산출하고, 상기 F1-score가 가장 높은 네트워크 특성 분류 모델을 상기 네트워크 특성 추출 파일의 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델로 결정하는, 네트워크

공격 탐지 장치.

청구항 6

청구항 5에 있어서,

상기 모델 결정부는,

상기 기 설정된 임계 값이 복수 개인 경우, 상기 복수의 임계 값 중 상기 F1-score가 가장 높을 때의 임계 값을 상기 네트워크 특성 추출 파일의 정상-비정상 분류의 기준이 될 임계 값으로 선택하는, 네트워크 공격 탐지 장치.

청구항 7

청구항 1에 있어서,

상기 공격 탐지부는,

기 설정된 종류의 네트워크 특성 각각에 대해 상기 결정된 네트워크 특성 분류 모델이 판단한 정상 특성 값 각각과 상기 네트워크 특성 추출 파일의 특성 값 각각에 기초하여 손실(loss)을 산출하고, 상기 손실이 기 설정된 임계 값 이상인 경우 상기 네트워크 특성 추출 파일을 비정상으로 분류하고, 상기 비정상으로 분류된 네트워크 특성 추출 파일에 대응되는 네트워크가 공격받은 것으로 탐지하는, 네트워크 공격 탐지 장치.

청구항 8

청구항 7에 있어서,

상기 공격 탐지부는,

상기 정상 특성 값 각각과 상기 네트워크 특성 추출 파일의 특성 값 각각 사이의 평균 제곱 오차(MSE; Mean Squared Error)를 상기 손실로 산출하는, 네트워크 공격 탐지 장치.

청구항 9

청구항 7에 있어서,

상기 공격 탐지부는,

기 설정된 제1 임계 값 및 상기 제1 임계 값을 초과하는 기 설정된 제2 임계 값에 있어서,

상기 손실이 상기 제1 임계 값 이상이고 상기 제2 임계 값 미만인 경우, 상기 네트워크 특성 추출 파일을 정상이되 위험군(risk group)으로 분류하고,

상기 손실이 상기 제2 임계 값 이상인 경우, 상기 네트워크 특성 추출 파일을 비정상으로 분류하는, 네트워크 공격 탐지 장치.

청구항 10

각각 복수의 네트워크 데이터 셋 중 하나와 복수의 딥 러닝(Deep learning) 알고리즘 중 하나를 이용한 학습을 통해 생성된 복수의 네트워크 특성 분류 모델을 저장하는 단계;

상기 복수의 네트워크 특성 분류 모델 중 분류 대상인 네트워크 특성 추출 파일에 대한 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정하는 단계;

상기 결정된 네트워크 특성 분류 모델을 이용하여 상기 네트워크 특성 추출 파일에 대해 상기 정상-비정상 분류

를 수행하는 단계; 및

상기 정상-비정상 분류의 결과에 기초하여 상기 네트워크 특성 추출 파일에 대응되는 네트워크에 대한 공격을 탐지하는 단계를 포함하되,

상기 결정하는 단계는,

정상-비정상 레이블(label)로 라벨링(labeling)된 복수의 테스트 파일을 포함하는 테스트 데이터 셋 및 상기 정상-비정상 분류의 기준이 되는 기 설정된 임계 값에 기초하여 상기 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정하는, 네트워크 공격 탐지 방법.

청구항 11

청구항 10에 있어서,

상기 복수의 딥 러닝 알고리즘은,

오토 인코더(Auto Encoder) 알고리즘 및 순환 신경망(RNN; Recurrent Neural Network) 알고리즘 중 적어도 하나를 포함하는, 네트워크 공격 탐지 방법.

청구항 12

청구항 10에 있어서,

네트워크 공격 탐지 장치에 입력되거나 상기 복수의 네트워크 데이터 셋 중 어느 하나에 포함된 복수의 네트워크 패킷 캡쳐 파일(network packet capture file)을 하나 이상의 네트워크 패킷 플로우(network packet flow)로 군집화하는 단계; 및

기 설정된 종류의 네트워크 특성에 기초하여 상기 하나 이상의 네트워크 패킷 플로우 각각에 대응되는 네트워크 특성 추출 파일을 생성하는 단계를 더 포함하는, 네트워크 공격 탐지 방법.

청구항 13

삭제

청구항 14

청구항 10에 있어서,

상기 결정하는 단계는,

상기 임계 값에 기초한 상기 복수의 테스트 파일의 정상-비정상 분류 결과와 상기 정상-비정상 레이블에 기초하여 상기 복수의 네트워크 특성 분류 모델 각각의 정밀도(Precision) 및 재현율(Recall)을 산출하는 단계;

상기 정밀도 및 상기 재현율에 기초하여 F1-score를 산출하는 단계; 및

상기 F1-score가 가장 높은 네트워크 특성 분류 모델을 상기 네트워크 특성 추출 파일의 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델로 결정하는 단계를 포함하는, 네트워크 공격 탐지 방법.

청구항 15

청구항 14에 있어서,

상기 결정하는 단계는,

상기 기 설정된 임계 값이 복수 개인 경우, 상기 복수의 임계 값 중 상기 F1-score가 가장 높을 때의 임계 값을 상기 네트워크 특성 추출 파일의 정상-비정상 분류의 기준이 될 임계 값으로 선택하는, 네트워크 공

격 탐지 방법.

청구항 16

청구항 10에 있어서,

상기 수행하는 단계는,

기 설정된 종류의 네트워크 특성 각각에 대해 상기 결정된 네트워크 특성 분류 모델이 판단한 정상 특성 값 각각과 상기 네트워크 특성 추출 파일의 특성 값 각각에 기초하여 손실(loss)을 산출하는 단계; 및

상기 손실이 기 설정된 임계 값 이상인 경우 상기 네트워크 특성 추출 파일을 비정상으로 분류하는 단계를 포함하고,

상기 탐지하는 단계는,

상기 비정상으로 분류된 네트워크 특성 추출 파일에 대응되는 네트워크가 공격받은 것으로 탐지하는, 네트워크 공격 탐지 방법.

청구항 17

청구항 16에 있어서,

상기 산출하는 단계는,

상기 정상 특성 값 각각과 상기 네트워크 특성 추출 파일의 특성 값 각각 사이의 평균 제곱 오차(MSE; Mean Squared Error)를 상기 손실로 산출하는, 네트워크 공격 탐지 방법.

청구항 18

청구항 16에 있어서,

상기 분류하는 단계는,

기 설정된 제1 임계 값 및 상기 제1 임계 값을 초과하는 기 설정된 제2 임계 값에 있어서,

상기 손실이 상기 제1 임계 값 이상이고 상기 제2 임계 값 미만인 경우, 상기 네트워크 특성 추출 파일을 정상이되 위험군(risk group)으로 분류하고,

상기 손실이 상기 제2 임계 값 이상인 경우, 상기 네트워크 특성 추출 파일을 비정상으로 분류하는, 네트워크 공격 탐지 방법.

발명의 설명

기 술 분 야

[0001] 개시되는 실시예들은 네트워크 공격을 탐지하는 기술에 관한 것이다.

배 경 기 술

[0003] 정보의 홍수 속에서 살아가는 우리는, 다양한 컴퓨팅 장치들이 상호 연결된 네트워크를 통해 수많은 데이터를 주고받으며 각자에게 필요한 정보를 습득한다. 따라서 원활한 정보 습득 및 공동체의 번영을 위해서는 안전한 환경의 네트워크를 구축하는 것이 중요한 과제라 할 수 있다.

[0004] 이를 위해서는 네트워크 교란, 정보 유출 등을 목적으로 하는 네트워크 공격을 정확하게 탐지할 필요가 있어, 종래에는 규칙(rule) 기반 탐지 방법과 머신 러닝(machine learning) 알고리즘 기반 탐지 방법이 활용되어왔다.

[0005] 그러나, 네트워크 공격이 점차 지능적이고 고도화됨에 따라 새로운 유형의 네트워크 공격이 등장하게 되었고, 종래의 규칙 기반 탐지 방법은 새로운 유형의 네트워크 공격에 대해서는 낮은 탐지율을 보일 뿐만 아니라 새로운 네트워크 공격 유형에 대한 규칙을 추가하기 위해서는 추가적으로 비용 및 시간이 소요된다는 문제가 있었다.

[0006] 한편, 종래의 머신 러닝 알고리즘 기반 탐지 방법은 새로운 유형의 네트워크 공격을 탐지할 수는 있으나, 이를 위한 학습(training)에는 비정상 네트워크와 관련된 상당한 양의 학습 데이터가 필요하다는 한계가 있으며, 특히 지도 학습(supervised learning) 기법을 통해 학습될 경우 데이터의 라벨링(labeling)에 추가적인 비용 및 시간이 소요된다는 문제가 있었다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 대한민국 등록특허공보 제10-0628325호 (2006.09.19. 등록)

발명의 내용

해결하려는 과제

[0009] 개시되는 실시예들은 비용, 시간, 데이터 셋의 구성 등의 제약을 극복하고 다양한 유형의 네트워크 공격을 탐지하기 위한 것이다.

과제의 해결 수단

[0011] 개시되는 일 실시예에 따른 네트워크 공격 탐지 장치는, 각각 복수의 정상 네트워크 데이터 셋 중 하나와 복수의 딥 러닝(Deep learning) 알고리즘 중 하나를 이용한 학습을 통해 생성된 복수의 네트워크 특성 분류 모델을 저장하는 모델 저장부, 상기 복수의 네트워크 특성 분류 모델 중 분류 대상인 네트워크 특성 추출 파일에 대한 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정하는 모델 결정부 및 상기 결정된 네트워크 특성 분류 모델을 이용하여 상기 네트워크 특성 추출 파일에 대해 상기 정상-비정상 분류를 수행하고, 상기 정상-비정상 분류의 결과에 기초하여 상기 네트워크 특성 추출 파일에 대응되는 네트워크에 대한 공격을 탐지하는 공격 탐지부를 포함한다.

[0012] 상기 복수의 딥 러닝 알고리즘은, 오토 인코더(Auto Encoder) 알고리즘 및 순환 신경망(RNN; Recurrent Neural Network) 알고리즘 중 적어도 하나를 포함할 수 있다.

[0013] 추가적인 실시예에 따른 네트워크 공격 탐지 장치는, 상기 공격 탐지부에 입력되거나 상기 복수의 정상 네트워크 데이터 셋 중 어느 하나에 포함된 복수의 네트워크 패킷 캡쳐 파일(network packet capture file)을 하나 이상의 네트워크 패킷 플로우(network packet flow)로 군집화하고, 기 설정된 종류의 네트워크 특성에 기초하여 상기 하나 이상의 네트워크 패킷 플로우 각각에 대응되는 네트워크 특성 추출 파일을 생성하는 전처리부를 더 포함할 수 있다.

[0014] 상기 모델 결정부는, 정상-비정상 레이블(label)로 라벨링(labeling)된 복수의 테스트 파일을 포함하는 테스트 데이터 셋 및 상기 정상-비정상 분류의 기준이 되는 기 설정된 임계 값에 기초하여 상기 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정할 수 있다.

[0015] 상기 모델 결정부는, 상기 임계 값에 기초한 상기 복수의 테스트 파일의 정상-비정상 분류 결과와 상기 정상-비정상 레이블에 기초하여 상기 복수의 네트워크 특성 분류 모델 각각의 정밀도(Precision) 및 재현율(Recall)을 산출하고, 상기 정밀도 및 상기 재현율에 기초하여 F1-score를 산출하고, 상기 F1-score가 가장 높은 네트워크 특성 분류 모델을 상기 네트워크 특성 추출 파일의 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델로 결정할 수 있다.

[0016] 상기 모델 결정부는, 상기 기 설정된 임계 값이 복수 개인 경우, 상기 복수의 임계 값 중 상기 F1-score가 가장 높을 때의 임계 값을 상기 네트워크 특성 추출 파일의 정상-비정상 분류의 기준이 될 임계 값으로 선택할 수 있다.

[0017] 상기 공격 탐지부는, 기 설정된 종류의 네트워크 특성 각각에 대해 상기 결정된 네트워크 특성 분류 모델이 판단한 정상 특성 값 각각과 상기 네트워크 특성 추출 파일의 특성 값 각각에 기초하여 손실(loss)을 산출하고, 상기 손실이 기 설정된 임계 값 이상인 경우 상기 네트워크 특성 추출 파일을 비정상으로 분류하고, 상기 비정상으로 분류된 네트워크 특성 추출 파일에 대응되는 네트워크가 공격받은 것으로 탐지할 수 있다.

[0018] 상기 공격 탐지부는, 상기 정상 특성 값 각각과 상기 네트워크 특성 추출 파일의 특성 값 각각 사이의 평균 제곱 오차(MSE; Mean Squared Error)를 상기 손실로 산출할 수 있다.

[0019] 상기 공격 탐지부는, 기 설정된 제1 임계 값 및 상기 제1 임계 값을 초과하는 기 설정된 제2 임계 값에 있어서, 상기 손실이 상기 제1 임계 값 이상이고 상기 제2 임계 값 미만인 경우, 상기 네트워크 특성 추출 파일을 정상이되 위험군(risk group)으로 분류할 수 있고, 상기 손실이 상기 제2 임계 값 이상인 경우, 상기 네트워크 특성 추출 파일을 비정상으로 분류할 수 있다.

[0020] 개시되는 일 실시예에 따른 네트워크 공격 탐지 방법은, 각각 복수의 네트워크 데이터 셋 중 하나와 복수의 딥 러닝(Deep learning) 알고리즘 중 하나를 이용한 학습을 통해 생성된 복수의 네트워크 특성 분류 모델을 저장하는 단계, 상기 복수의 네트워크 특성 분류 모델 중 분류 대상인 네트워크 특성 추출 파일에 대한 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정하는 단계, 상기 결정된 네트워크 특성 분류 모델을 이용하여 상기 네트워크 특성 추출 파일에 대해 상기 정상-비정상 분류를 수행하는 단계 및 상기 정상-비정상 분류의 결과에 기초하여 상기 네트워크 특성 추출 파일에 대응되는 네트워크에 대한 공격을 탐지하는 단계를 포함한다.

[0021] 상기 복수의 딥 러닝 알고리즘은, 오토 인코더(Auto Encoder) 알고리즘 및 순환 신경망(RNN; Recurrent Neural Network) 알고리즘 중 적어도 하나를 포함할 수 있다.

[0022] 추가적인 실시예에 따른 네트워크 공격 탐지 방법은, 네트워크 공격 탐지 장치에 입력되거나 상기 복수의 네트워크 데이터 셋 중 어느 하나에 포함된 복수의 네트워크 패킷 캡쳐 파일(network packet capture file)을 하나 이상의 네트워크 패킷 플로우(network packet flow)로 군집화하는 단계 및 기 설정된 종류의 네트워크 특성에 기초하여 상기 하나 이상의 네트워크 패킷 플로우 각각에 대응되는 네트워크 특성 추출 파일을 생성하는 단계를 더 포함할 수 있다.

[0023] 상기 결정하는 단계는, 정상-비정상 레이블(label)로 라벨링(labeling)된 복수의 테스트 파일을 포함하는 테스트 데이터 셋 및 상기 정상-비정상 분류의 기준이 되는 기 설정된 임계 값에 기초하여 상기 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정할 수 있다.

[0024] 상기 결정하는 단계는, 상기 임계 값에 기초한 상기 복수의 테스트 파일의 정상-비정상 분류 결과와 상기 정상-비정상 레이블에 기초하여 상기 복수의 네트워크 특성 분류 모델 각각의 정밀도(Precision) 및 재현율(Recall)을 산출하는 단계, 상기 정밀도 및 상기 재현율에 기초하여 F1-score를 산출하는 단계 및 상기 F1-score가 가장 높은 네트워크 특성 분류 모델을 상기 네트워크 특성 추출 파일의 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델로 결정하는 단계를 포함할 수 있다.

[0025] 상기 결정하는 단계는, 상기 기 설정된 임계 값이 복수 개인 경우, 상기 복수의 임계 값 중 상기 F1-score가 가장 높을 때의 임계 값을 상기 네트워크 특성 추출 파일의 정상-비정상 분류의 기준이 될 임계 값으로 선택할 수 있다.

[0026] 상기 수행하는 단계는, 기 설정된 종류의 네트워크 특성 각각에 대해 상기 결정된 네트워크 특성 분류 모델이 판단한 정상 특성 값 각각과 상기 네트워크 특성 추출 파일의 특성 값 각각에 기초하여 손실(loss)을 산출하는 단계 및 상기 손실이 기 설정된 임계 값 이상인 경우 상기 네트워크 특성 추출 파일을 비정상으로 분류하는 단계를 포함할 수 있고, 상기 탐지하는 단계는, 상기 비정상으로 분류된 네트워크 특성 추출 파일에 대응되는 네트워크가 공격받은 것으로 탐지할 수 있다.

[0027] 상기 산출하는 단계는, 상기 정상 특성 값 각각과 상기 네트워크 특성 추출 파일의 특성 값 각각 사이의 평균 제곱 오차(MSE; Mean Squared Error)를 상기 손실로 산출할 수 있다.

[0028] 상기 분류하는 단계는, 기 설정된 제1 임계 값 및 상기 제1 임계 값을 초과하는 기 설정된 제2 임계 값에 있어서, 상기 손실이 상기 제1 임계 값 이상이고 상기 제2 임계 값 미만인 경우, 상기 네트워크 특성 추출 파일을 정상이되 위험군(risk group)으로 분류할 수 있고, 상기 손실이 상기 제2 임계 값 이상인 경우, 상기 네트워크 특성 추출 파일을 비정상으로 분류할 수 있다.

발명의 효과

[0030] 개시되는 실시예들에 따르면, 딥 러닝(Deep learning) 알고리즘에 기반한 학습을 통해 생성된 모델로 네트워크에 대한 공격을 탐지함으로써, 새로운 유형의 네트워크 공격을 탐지하기 위해 필요한 학습 데이터의 양, 비용 및 시간을 절감할 수 있다.

[0031] 또한 개시되는 실시예들에 따르면, 다양한 딥 러닝 알고리즘에 기반하여 학습된 여러 모델 중 정밀도와 재현율을 고려하여 가장 우수한 모델을 사용함으로써, 현존하는 다양한 딥 러닝 알고리즘 및 향후 개선될 알고리즘들을 용이하게 모델의 학습에 사용할 수 있다.

도면의 간단한 설명

[0033] 도 1은 일 실시예에 따른 네트워크 공격 탐지 장치를 설명하기 위한 블록도

도 2는 추가적인 실시예에 따른 네트워크 공격 탐지 장치를 설명하기 위한 블록도

도 3은 일 실시예에 따른 네트워크 공격 탐지 방법을 설명하기 위한 흐름도

도 4는 추가적인 실시예에 따른 네트워크 공격 탐지 방법을 설명하기 위한 흐름도

도 5는 일 실시예에 따른 320 단계를 보다 상세히 설명하기 위한 흐름도

도 6은 일 실시예에 따른 330 단계를 보다 상세히 설명하기 위한 흐름도

도 7은 추가적인 실시예에 따른 330 단계를 보다 상세히 설명하기 위한 흐름도

도 8은 일 실시예에 따른 컴퓨팅 장치를 포함하는 컴퓨팅 환경을 예시하여 설명하기 위한 블록도

발명을 실시하기 위한 구체적인 내용

[0034] 이하, 도면을 참조하여 구체적인 실시형태를 설명하기로 한다. 이하의 상세한 설명은 본 명세서에서 기술된 방법, 장치 및/또는 시스템에 대한 포괄적인 이해를 돕기 위해 제공된다. 그러나 이는 예시에 불과하며 개시되는 실시예들은 이에 제한되지 않는다.

[0035] 실시예들을 설명함에 있어서, 관련된 공지기술에 대한 구체적인 설명이 개시되는 실시예들의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하기로 한다. 그리고, 후술되는 용어들은 개시되는 실시예들에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다. 상세한 설명에서 사용되는 용어는 단지 실시예들을 기술하기 위한 것이며, 결코 제한적이어서는 안 된다. 명확하게 달리 사용되지 않는 한, 단수 형태의 표현은 복수 형태의 의미를 포함한다. 본 설명에서, "포함" 또는 "구비"와 같은 표현은 어떤 특성들, 숫자들, 단계들, 동작들, 요소들, 이들의 일부 또는 조합을 가리키기 위한 것이며, 기술된 것 이외에 하나 또는 그 이상의 다른 특성, 숫자, 단계, 동작, 요소, 이들의 일부 또는 조합의 존재 또는 가능성을 배제하도록 해석되어서는 안 된다.

[0036] 도 1은 일 실시예에 따른 네트워크 공격 탐지 장치(100)를 설명하기 위한 블록도이다.

[0037] 이하의 실시예들에서, '네트워크'는 인터넷, 하나 이상의 로컬 영역 네트워크(local area networks), 광역 네트워크(wire area networks), 셀룰러 네트워크, 모바일 네트워크, 그 밖에 다른 종류의 네트워크들, 또는 이러한 네트워크들의 조합을 포함할 수 있다.

[0038] 한편 이하의 실시예들에서, '네트워크 공격'은 네트워크 상의 정보, 네트워크에 연결된 기기에 저장된 정보 또는 네트워크 자체를 교란, 거부, 손상, 파괴하는 행위를 포함할 수 있다. 예를 들어, '네트워크 공격'은 '시스템 거부공격(Denial of Service), 스니핑(Sniffing), 스푸핑(Spoofing), 스위치 재밍(Switch Jamming), 랜드 어택(Land Attack) 또는 포트 스캐닝(PORT Scanning) 등을 포함할 수 있으나, 반드시 이에 한정되는 것은 아니다.

[0039] 도시된 바와 같이, 일 실시에 따른 네트워크 공격 탐지 장치(100)는 모델 저장부(110), 모델 결정부(120) 및 공격 탐지부(130)를 포함한다.

[0040] 모델 저장부(110)는 각각 복수의 정상 네트워크 데이터 셋 중 하나와 복수의 딥 러닝(Deep learning) 알고리즘

중 하나를 이용한 학습을 통해 생성된 복수의 네트워크 특성 분류 모델을 저장한다.

[0041] 일 실시예에 따르면, 복수의 정상 네트워크 데이터 셋 각각은 정상 네트워크에 대한 복수의 네트워크 패킷 캡쳐 파일(network packet capture file)을 포함할 수 있다. 또한, 일 실시예에 따르면, 복수의 정상 네트워크 데이터 셋 각각은 정상 네트워크에 대한 복수의 네트워크 특성 추출 파일을 포함할 수도 있다.

[0042] 이하의 실시예들에서, '네트워크 패킷'은 네트워크가 전달하는 네트워크 데이터의 형식화된 블록(block)을 의미하며, '네트워크 패킷 캡쳐 파일'은 이러한 네트워크 패킷을 저장하기 위해 사용되는 파일을 의미한다.

[0043] 한편, 이하의 실시예들에서, '네트워크 특성'은 네트워크에 대한 공격 여부를 탐지하는 기준이 복수의 기준을 의미한다. 예를 들어, '네트워크 특성'은 '통신한 총 네트워크 패킷의 개수', '전체 통신 시간', '네트워크 패킷의 통신에 소요된 시간의 표준편차', '네트워크 패킷의 통신에 소요된 시간의 평균', '네트워크 패킷의 통신에 소요된 시간의 최댓값', '네트워크 패킷의 통신에 소요된 시간의 최솟값', '1초당 통신한 네트워크 패킷의 개수', 'TCP 송신 포트(Source Port)와 수신 포트(Destination Port)의 비율', '네트워크 패킷의 헤더(header)에 사용된 전체 바이트(bite)', '최초 윈도우(window)에서 보내진 바이트', '네트워크 패킷의 최대 사이즈(size)', '네트워크 패킷의 최소 사이즈' 등을 포함할 수 있으나, 반드시 이에 한정되는 것은 아니며, 네트워크에 대한 공격 여부를 탐지하는 데 사용될 수 있는 임의의 기준을 더 포함할 수도 있다.

[0044] 일 실시예에 따르면, 복수의 딥 러닝 알고리즘은, 오토 인코더(Auto Encoder) 알고리즘 및 순환 신경망(RNN; Recurrent Neural Network) 알고리즘 중 적어도 하나를 포함할 수 있다.

[0045] 구체적으로, 복수의 딥 러닝 알고리즘은, RNN 알고리즘 중에서 LSTM(Long Short-Term Memory) 알고리즘을 포함할 수 있으나, 반드시 이에 한정되는 것은 아니다.

[0046] 모델 결정부(120)는 복수의 네트워크 특성 분류 모델 중 분류 대상인 네트워크 특성 추출 파일에 대한 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정한다.

[0047] 일 실시예에 따르면, 모델 결정부(120)는 정상-비정상 레이블(label)로 라벨링(labeling)된 복수의 테스트 파일을 포함하는 테스트 데이터 셋 및 정상-비정상 분류의 기준이 되는 기 설정된 임계 값에 기초하여 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정할 수 있다.

[0048] 이때, 테스트 데이터 셋에 포함된 복수의 테스트 파일은 기 설정된 종류의 네트워크 특성 각각을 기준으로 정상-비정상이 구분되어 라벨링된 복수의 네트워크 특성 추출 파일일 수 있다.

[0049] 구체적으로, 모델 결정부(120)는 아래의 과정을 통해 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정할 수 있다.

[0050] (1) 기 설정된 임계 값에 기초한 복수의 테스트 파일의 정상-비정상 분류 결과와 정상-비정상 레이블에 기초하여, 복수의 네트워크 특성 분류 모델 각각의 정밀도(Precision) 및 재현율(Recall)을 산출

[0051] (2) 정밀도 및 재현율에 기초하여 F1-score를 산출

[0052] (3) F1-score가 가장 높은 네트워크 특성 분류 모델을 네트워크 특성 추출 파일의 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델로 결정

[0053] 이하의 실시예들에서, '정밀도'는 복수의 테스트 파일을 복수의 네트워크 특성 분류 모델을 통해 각각 정상 또는 비정상으로 분류한 결과, 비정상으로 분류된 테스트 파일 중 실제 비정상으로 라벨링된 테스트 파일의 비율을 의미할 수 있다.

[0054] 즉 다시 말하면, '정밀도'는 아래의 수학식 1에 의해 산출될 수 있다.

[0055] [수학식 1]

$$정밀도(Precision) = \frac{TP}{TP + FP}$$

[0056]

[0057] 이때, TP(True Positive)는 비정상으로 분류되고 실제 비정상으로 라벨링된 테스트 파일의 수, FP(False Positive)는 비정상으로 분류되나 실제 정상으로 라벨링된 테스트 파일의 수를 나타낸다.

[0058] 한편 이하의 실시예들에서, '재현율'은 복수의 테스트 파일을 복수의 네트워크 특성 분류 모델을 통해 각각 정

상 또는 비정상으로 분류한 결과, 실제 비정상으로 라벨링된 테스트 파일 중에서 비정상으로 분류된 테스트 파일의 비율을 의미할 수 있다.

[0059] 즉 다시 말하면, '재현율'은 아래의 수학식 2에 의해 산출될 수 있다.

[0060] [수학식 2]

$$재현율(Recall) = \frac{TP}{TP + FN}$$

[0061]

[0062] 이때, FN(False Negative)은 정상으로 분류되나 실제 비정상으로 라벨링된 테스트 파일의 수를 나타낸다.

[0063] 일 실시예에 따르면, 모델 결정부(120)는 정밀도와 재현율의 조화평균(Harmonic Mean)을 F1-score로 산출할 수 있다.

[0064] 구체적으로, F1-score는 아래의 수학식 3에 의해 산출될 수 있다.

[0065] [수학식 3]

$$조화평균(Harmonic\ Mean) = \frac{2 * 정밀도 * 재현율}{정밀도 + 재현율}$$

[0066]

[0067] 일 실시예에 따르면, 모델 결정부(120)는 기 설정된 임계 값이 복수 개인 경우, 복수의 기 설정된 임계 값 중 F1-score가 가장 높을 때의 임계 값을 네트워크 특성 추출 파일의 정상-비정상 분류의 기준이 될 임계 값으로 선택할 수 있다.

[0068] 즉, 테스트 데이터 셋에 포함된 테스트 파일을 이용하여 임계 값을 설정하는 경우, 정상 네트워크 데이터 셋으로부터 획득한 네트워크 특성 추출 파일을 이용하여 임계 값을 설정할 때와 비교하여 정상-비정상 분류의 신뢰도에 차이가 발생할 수 있다. 이를 보완하기 위해, 네트워크 공격 탐지 장치(100)는 임계 값을 복수 개 설정하고, 모델 결정부(120)로 하여금 F1-score가 가장 높을 때의 임계 값을 선택하도록 할 수 있다.

[0069] 예를 들어, 정상 네트워크 데이터 셋으로부터 획득한 네트워크 특성 추출 파일을 이용하여 설정한 신뢰도 95%의 임계 값이 0.95라 가정하자. 이 경우, 네트워크 공격 탐지 장치(100)는 0.95에서 0.01 또는 0.02를 더하거나 뺀 값인 0.93, 0.94, 0.96, 0.97을 추가적인 임계 값으로 설정하여, 모델 결정부(120)로 하여금 총 5개의 임계 값 중 F1-score가 가장 높을 때의 임계 값을 선택하도록 할 수 있다.

[0070] 공격 탐지부(130)는 결정된 네트워크 특성 분류 모델을 이용하여 네트워크 특성 추출 파일에 대해 정상-비정상 분류를 수행하고, 정상-비정상 분류의 결과에 기초하여 네트워크 특성 추출 파일에 대응되는 네트워크에 대한 공격을 탐지한다.

[0071] 일 실시예에 따르면, 공격 탐지부(130)는 아래의 과정을 통해 네트워크에 대한 공격을 탐지할 수 있다.

[0072] (1) 기 설정된 종류의 네트워크 특성 각각에 대해, 모델 결정부(120)를 통해 결정된 네트워크 특성 분류 모델이 판단한 정상 특성 값 각각과 네트워크 특성 추출 파일의 특성 값 각각에 기초하여 손실(loss)을 산출

[0073] (2) 산출된 손실이 기 설정된 임계 값 이상인 경우, 네트워크 특성 추출 파일을 비정상으로 분류

[0074] (3) 비정상으로 분류된 네트워크 특성 추출 파일에 대응되는 네트워크가 공격받은 것으로 탐지

[0075] 구체적으로, 공격 탐지부(130)는 정상 특성 값 각각과 네트워크 특성 추출 파일의 특성 값 각각 사이의 평균 제곱 오차(MSE; Mean Squared Error)를 손실로 산출할 수 있다.

[0076] 이와 관련하여, 정상-비정상 분류의 기준이 되는 임계 값은 정상 특성 값 각각과 정상 네트워크 데이터 셋으로부터 획득한 네트워크 특성 추출 파일의 특성 값 각각 사이의 MSE들로 이루어진 분포에서, 상위 X 퍼센티지(이때, X는 0 이상 100 이하의 양의 실수)에 해당하는 MSE일 수 있다.

[0077] 예를 들어, 기 설정되는 임계 값은 정상 특성 값 각각과 정상 네트워크 데이터 셋으로부터 획득한 네트워크 특성 추출 파일의 특성 값 각각 사이의 MSE들로 이루어진 분포에서, 최대값에 해당하는 MSE일 수 있다.

[0078] 일 실시예에 따르면, 공격 탐지부(130)는 기 설정된 제1 임계 값 및 제1 임계 값을 초과하는 기 설정된 제2 임

계 값에 있어서, 손실이 제1 임계 값 이상이고 제2 임계 값 미만인 경우, 네트워크 특성 추출 파일을 정상이되 위험군(risk group)으로 분류할 수 있다.

[0079] 한편, 일 실시예에 따르면, 공격 탐지부(130)는 손실이 제2 임계 값 이상인 경우, 네트워크 특성 추출 파일을 비정상으로 분류할 수 있다.

[0080] 또한, 일 실시예에 따르면, 공격 탐지부(130)는 손실이 제1 임계 값 미만인 경우, 네트워크 특성 추출 파일을 정상이되 비위험군(non-risk group)으로 분류할 수 있다.

[0081] 예를 들어, 제1 임계 값은 정상 특성 값 각각과 정상 네트워크 데이터 셋으로부터 획득한 네트워크 특성 추출 파일의 특성 값 각각 사이의 MSE들로 이루어진 분포에서, 상위 5퍼센티지에 해당하는 MSE일 수 있으며, 제2 임계 값은 최대값에 해당하는 MSE일 수 있다.

[0082] 도 2는 추가적인 실시예에 따른 네트워크 공격 탐지 장치(200)를 설명하기 위한 블록도이다.

[0083] 도시된 바와 같이, 추가적인 실시예에 따른 네트워크 공격 탐지 장치(200)는 모델 저장부(110), 모델 결정부(120) 및 공격 탐지부(130) 외에 전처리부(210)를 더 포함한다.

[0084] 이 중, 모델 저장부(110), 모델 결정부(120) 및 공격 탐지부(130)는 도 1을 참조하여 설명한 일 실시예에서와 동일 또는 유사한 기능을 수행하므로, 이에 대한 중복되는 설명은 생략하도록 한다.

[0085] 전처리부(210)는 공격 탐지부에 입력되거나 복수의 정상 네트워크 데이터 셋 중 어느 하나에 포함된 복수의 네트워크 패킷 캡쳐 파일을 하나 이상의 네트워크 패킷 플로우(network packet flow)로 군집화할 수 있다.

[0086] 이하의 실시예들에서, '네트워크 패킷 플로우'는 송-수신 IP 및 송-수신 포트 번호 중 적어도 하나가 동일한 네트워크 패킷들에 대해, 해당 네트워크 패킷들에 대응되는 복수의 네트워크 패킷 캡쳐 파일을 한데 묶은 군집일 수 있다. 그러나, 네트워크 패킷 캡쳐 파일을 분류하는 기준은 이외에도 다양할 수 있으며, 해당 기준을 달리함에 따라 네트워크 패킷 플로우를 구성하는 네트워크 패킷 캡쳐 파일들의 구성이 달라질 수 있음은 자명하다.

[0087] 아울러, 전처리부(210)는 기 설정된 종류의 네트워크 특성에 기초하여, 군집화된 하나 이상의 네트워크 패킷 플로우 각각에 대응되는 네트워크 특성 추출 파일을 생성할 수 있다.

[0088] 도 3은 일 실시예에 따른 네트워크 공격 탐지 방법을 설명하기 위한 흐름도이다. 도 3에 도시된 방법은 예를 들어, 도 1을 참조하여 상술한 네트워크 공격 탐지 장치(100)에 의해 수행될 수 있다.

[0089] 우선, 네트워크 공격 탐지 장치(100)는 각각 복수의 네트워크 데이터 셋 중 하나와 복수의 딥 러닝(Deep learning) 알고리즘 중 하나를 이용한 학습을 통해 생성된 복수의 네트워크 특성 분류 모델을 저장한다(310).

[0090] 이후, 네트워크 공격 탐지 장치(100)는 복수의 네트워크 특성 분류 모델 중 분류 대상인 네트워크 특성 추출 파일에 대한 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정한다(320).

[0091] 이후, 네트워크 공격 탐지 장치(100)는 결정된 네트워크 특성 분류 모델을 이용하여 네트워크 특성 추출 파일에 대해 정상-비정상 분류를 수행한다(330).

[0092] 이후, 네트워크 공격 탐지 장치(100)는 정상-비정상 분류의 결과에 기초하여 네트워크 특성 추출 파일에 대응되는 네트워크에 대한 공격을 탐지한다(340).

[0093] 도 4는 추가적인 실시예에 따른 네트워크 공격 탐지 방법을 설명하기 위한 흐름도이다. 도 4에 도시된 방법은 예를 들어, 도 2를 참조하여 상술한 네트워크 공격 탐지 장치(200)에 의해 수행될 수 있다.

[0094] 우선, 네트워크 공격 탐지 장치(200)는 네트워크 공격 탐지 장치에 입력되거나 복수의 네트워크 데이터 셋 중 어느 하나에 포함된 복수의 네트워크 패킷 캡쳐 파일(network packet capture file)을 하나 이상의 네트워크 패킷 플로우(network packet flow)로 군집화한다(410).

[0095] 이후, 네트워크 공격 탐지 장치(200)는 기 설정된 종류의 네트워크 특성에 기초하여 하나 이상의 네트워크 패킷 플로우 각각에 대응되는 네트워크 특성 추출 파일을 생성한다(420).

[0096] 이후, 네트워크 공격 탐지 장치(200)는 각각 복수의 네트워크 데이터 셋 중 하나와 복수의 딥 러닝 알고리즘 중 하나를 이용한 학습을 통해 생성된 복수의 네트워크 특성 분류 모델을 저장한다(430).

[0097] 이후, 네트워크 공격 탐지 장치(200)는 복수의 네트워크 특성 분류 모델 중 분류 대상인 네트워크 특성 추출 파일에 대한 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정한다(440).

[0098] 이후, 네트워크 공격 탐지 장치(200)는 결정된 네트워크 특성 분류 모델을 이용하여 네트워크 특성 추출 파일에 대해 정상-비정상 분류를 수행한다(450).

[0099] 이후, 네트워크 공격 탐지 장치(200)는 정상-비정상 분류의 결과에 기초하여 네트워크 특성 추출 파일에 대응되는 네트워크에 대한 공격을 탐지한다(460).

[0100] 도 5는 일 실시예에 따른 320 단계를 보다 상세히 설명하기 위한 흐름도이다. 도 5에 도시된 방법은 예를 들어, 도 1을 참조하여 상술한 네트워크 공격 탐지 장치(100)에 의해 수행될 수 있다. 그러나, 320 단계는 도 4를 참조하여 설명한 440 단계와 대응되며, 이에 따라 도 5에 도시된 방법은 도 2를 참조하여 상술한 네트워크 공격 탐지 장치(200)에 의해서도 수행될 수 있다.

[0101] 우선, 네트워크 공격 탐지 장치(100, 200)는 임계 값에 기초한 복수의 테스트 파일의 정상-비정상 분류 결과와 정상-비정상 레이블에 기초하여 복수의 네트워크 특성 분류 모델 각각의 정밀도(Precision) 및 재현율(Recall)을 산출한다(510).

[0102] 이후, 네트워크 공격 탐지 장치(100, 200)는 정밀도 및 재현율에 기초하여 F1-score를 산출한다(520).

[0103] 이후, 네트워크 공격 탐지 장치(100, 200)는 F1-score가 가장 높은 네트워크 특성 분류 모델을 네트워크 특성 추출 파일의 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델로 결정한다(530).

[0104] 도 6은 일 실시예에 따른 330 단계를 보다 상세히 설명하기 위한 흐름도이다. 도 6에 도시된 방법은 예를 들어, 도 1을 참조하여 상술한 네트워크 공격 탐지 장치(100)에 의해 수행될 수 있다. 그러나, 330 단계는 도 4를 참조하여 설명한 450 단계와 대응되며, 이에 따라 도 6에 도시된 방법은 도 2를 참조하여 상술한 네트워크 공격 탐지 장치(200)에 의해서도 수행될 수 있다.

[0105] 우선, 네트워크 공격 탐지 장치(100, 200)는 기 설정된 종류의 네트워크 특성 각각에 대해 결정된 네트워크 특성 분류 모델이 판단한 정상 특성 값 각각과 네트워크 특성 추출 파일의 특성 값 각각에 기초하여 손실(loss)을 산출한다(610).

[0106] 이후, 네트워크 공격 탐지 장치(100, 200)는 산출된 손실이 기 설정된 임계 값 이상인지 여부를 판단한다(620).

[0107] 이후, 네트워크 공격 탐지 장치(100, 200)는 손실이 기 설정된 임계 값 이상인 경우, 네트워크 특성 추출 파일을 비정상으로 분류한다(630).

[0108] 한편, 네트워크 공격 탐지 장치(100, 200)는 손실이 기 설정된 임계 값 미만인 경우, 네트워크 특성 추출 파일을 정상으로 분류한다(640).

[0109] 도 7은 추가적인 실시예에 따른 330 단계를 보다 상세히 설명하기 위한 흐름도이다. 도 7에 도시된 방법은 예를 들어, 도 1을 참조하여 상술한 네트워크 공격 탐지 장치(100)에 의해 수행될 수 있다. 그러나, 330 단계는 도 4를 참조하여 설명한 450 단계와 대응되며, 이에 따라 도 7에 도시된 방법은 도 2를 참조하여 상술한 네트워크 공격 탐지 장치(200)에 의해서도 수행될 수 있다.

[0110] 우선, 네트워크 공격 탐지 장치(100, 200)는 기 설정된 종류의 네트워크 특성 각각에 대해 결정된 네트워크 특성 분류 모델이 판단한 정상 특성 값 각각과 네트워크 특성 추출 파일의 특성 값 각각에 기초하여 손실을 산출한다(710).

[0111] 이후, 네트워크 공격 탐지 장치(100, 200)는 산출된 손실이 기 설정된 제1 임계 값 이상인지 여부를 판단한다(720).

[0112] 이후, 네트워크 공격 탐지 장치(100, 200)는 손실이 제1 임계 값 이상인 경우, 제1 임계 값을 초과하는 기 설정된 제2 임계 값을 기준으로 하여, 손실이 제2 임계 값 이상인지 여부를 판단한다(730).

[0113] 이후, 네트워크 공격 탐지 장치(100, 200)는 손실이 제2 임계 값 이상인 경우, 네트워크 특성 추출 파일을 비정상으로 분류한다(740).

[0114] 한편, 네트워크 공격 탐지 장치(100, 200)는 손실이 제1 임계 값 이상이고 제2 임계 값 미만인 경우, 네트워크 특성 추출 파일을 정상이되 위험군(risk group)으로 분류한다(750).

[0115] 한편, 네트워크 공격 탐지 장치(100, 200)는 손실이 제1 임계 값 미만인 경우, 네트워크 특성 추출 파일을 정상이되 비위험군(non-risk group)으로 분류한다(760).

[0116] 이상의 흐름도 도 3 내지 도 7에서는 상기 방법을 복수 개의 단계로 나누어 기재하였으나, 적어도 일부의 단계들은 순서를 바꾸어 수행되거나, 다른 단계와 결합되어 함께 수행되거나, 생략되거나, 세부 단계들로 나뉘어 수행되거나, 또는 도시되지 않은 하나 이상의 단계가 부가되어 수행될 수 있다.

[0117] 도 8은 일 실시예에 따른 컴퓨팅 장치를 포함하는 컴퓨팅 환경(10)을 예시하여 설명하기 위한 블록도이다. 도시된 실시예에서, 각 컴포넌트들은 이하에 기술된 것 이외에 상이한 기능 및 능력을 가질 수 있고, 이하에 기술된 것 이외에도 추가적인 컴포넌트를 포함할 수 있다.

[0118] 도시된 컴퓨팅 환경(10)은 컴퓨팅 장치(12)를 포함한다. 일 실시예에서, 컴퓨팅 장치(12)는 일 실시예에 따른 네트워크 공격 탐지 장치(100)일 수 있다. 또한, 컴퓨팅 장치(12)는 추가적인 실시예에 따른 네트워크 공격 탐지 장치(200)일 수 있다.

[0119] 컴퓨팅 장치(12)는 적어도 하나의 프로세서(14), 컴퓨터 판독 가능 저장 매체(16) 및 통신 버스(18)를 포함한다. 프로세서(14)는 컴퓨팅 장치(12)로 하여금 앞서 언급된 예시적인 실시예에 따라 동작하도록 할 수 있다. 예컨대, 프로세서(14)는 컴퓨터 판독 가능 저장 매체(16)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 상기 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 상기 컴퓨터 실행 가능 명령어는 프로세서(14)에 의해 실행되는 경우 컴퓨팅 장치(12)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.

[0120] 컴퓨터 판독 가능 저장 매체(16)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능 저장 매체(16)에 저장된 프로그램(20)은 프로세서(14)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독 가능 저장 매체(16)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 컴퓨팅 장치(12)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.

[0121] 통신 버스(18)는 프로세서(14), 컴퓨터 판독 가능 저장 매체(16)를 포함하여 컴퓨팅 장치(12)의 다른 다양한 컴포넌트들을 상호 연결한다.

[0122] 컴퓨팅 장치(12)는 또한 하나 이상의 입출력 장치(24)를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(22) 및 하나 이상의 네트워크 통신 인터페이스(26)를 포함할 수 있다. 입출력 인터페이스(22) 및 네트워크 통신 인터페이스(26)는 통신 버스(18)에 연결된다. 입출력 장치(24)는 입출력 인터페이스(22)를 통해 컴퓨팅 장치(12)의 다른 컴포넌트들에 연결될 수 있다. 예시적인 입출력 장치(24)는 포인팅 장치(마우스 또는 트랙패드 등), 키보드, 터치 입력 장치(터치패드 또는 터치스크린 등), 음성 또는 소리 입력 장치, 다양한 종류의 센서 장치 및/또는 촬영 장치와 같은 입력 장치, 및/또는 디스플레이 장치, 프린터, 스피커 및/또는 네트워크 카드와 같은 출력 장치를 포함할 수 있다. 예시적인 입출력 장치(24)는 컴퓨팅 장치(12)를 구성하는 일 컴포넌트로서 컴퓨팅 장치(12)의 내부에 포함될 수도 있고, 컴퓨팅 장치(12)와는 구별되는 별개의 장치로 컴퓨팅 장치(12)와 연결될 수도 있다.

[0123] 한편, 본 발명의 실시예는 본 명세서에서 기술한 방법들을 컴퓨터상에서 수행하기 위한 프로그램, 및 상기 프로그램을 포함하는 컴퓨터 판독 가능 기록매체를 포함할 수 있다. 상기 컴퓨터 판독 가능 기록매체는 프로그램 명령, 로컬 데이터 파일, 로컬 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체는 본 발명을 위하여 특별히 설계되고 구성된 것들이거나, 또는 컴퓨터 소프트웨어 분야에서 통상적으로 사용 가능한 것일 수 있다. 컴퓨터 판독 가능 기록매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM, DVD와 같은 광 기록 매체, 및 롬, 램, 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 상기 프로그램의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함할 수 있다.

[0124] 이상에서 본 발명의 대표적인 실시예들을 상세하게 설명하였으나, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자는 상술한 실시예에 대하여 본 발명의 범주에서 벗어나지 않는 한도 내에서 다양한 변형이 가능함을 이해할 것이다. 그러므로 본 발명의 권리범위는 설명된 실시예에 국한되어 정해져서는 안 되며, 후술하는 청구범위뿐만 아니라 이 청구범위와 균등한 것들에 의해 정해져야 한다.

부호의 설명

[0126] 10: 컴퓨팅 환경

12: 컴퓨팅 장치

14: 프로세서

16: 컴퓨터 판독 가능 저장 매체

18: 통신 버스

20: 프로그램

22: 입출력 인터페이스

24: 입출력 장치

26: 네트워크 통신 인터페이스

100, 200: 네트워크 공격 탐지 장치

110: 모델 저장부

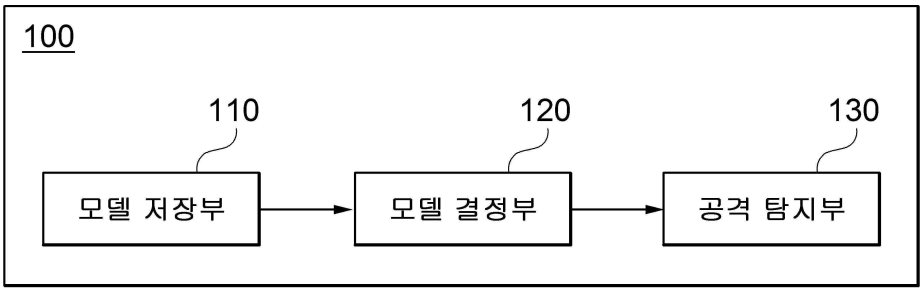
120: 모델 결정부

130: 공격 탐지부

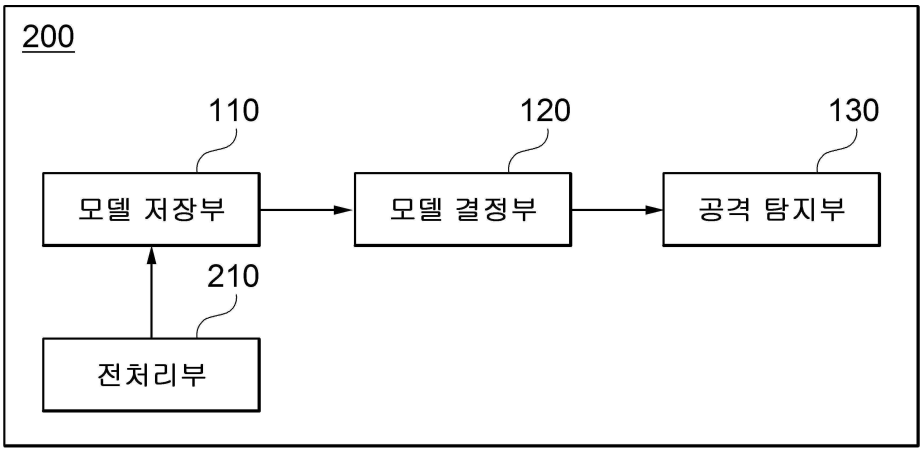
210: 전처리부

도면

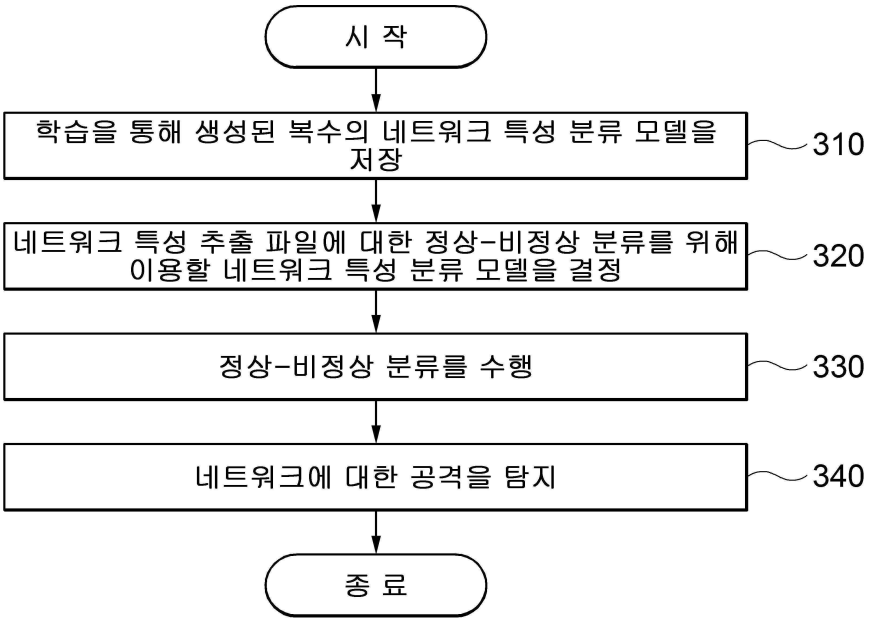
도면1



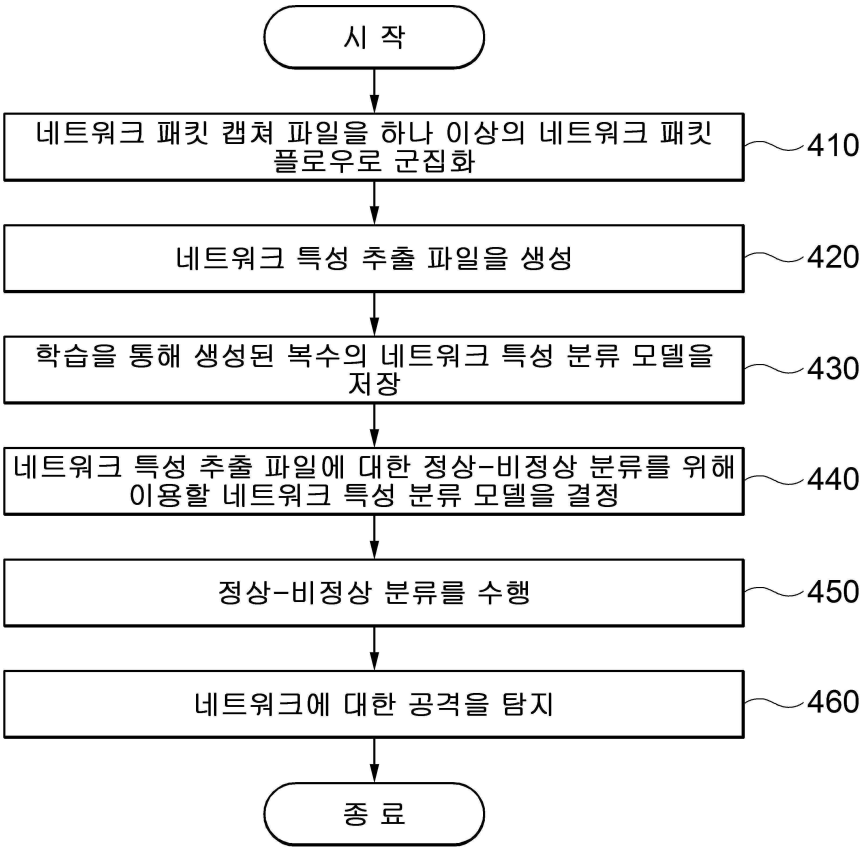
도면2



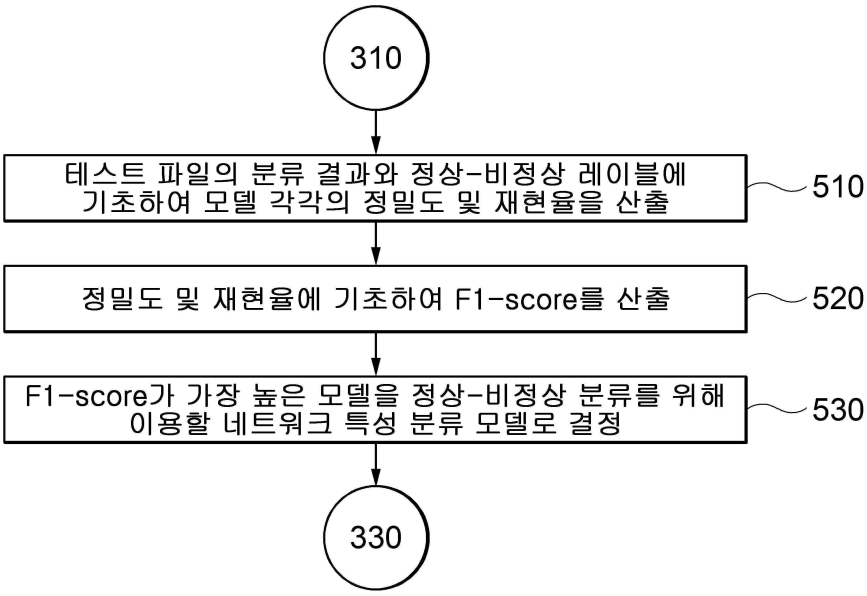
도면3


시 작
학습을 통해 생성된 복수의 네트워크 특성 분류 모델을 저장
310
네트워크 특성 추출 파일에 대한 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정
320
정상-비정상 분류를 수행
330
네트워크에 대한 공격을 탐지
340
종 료

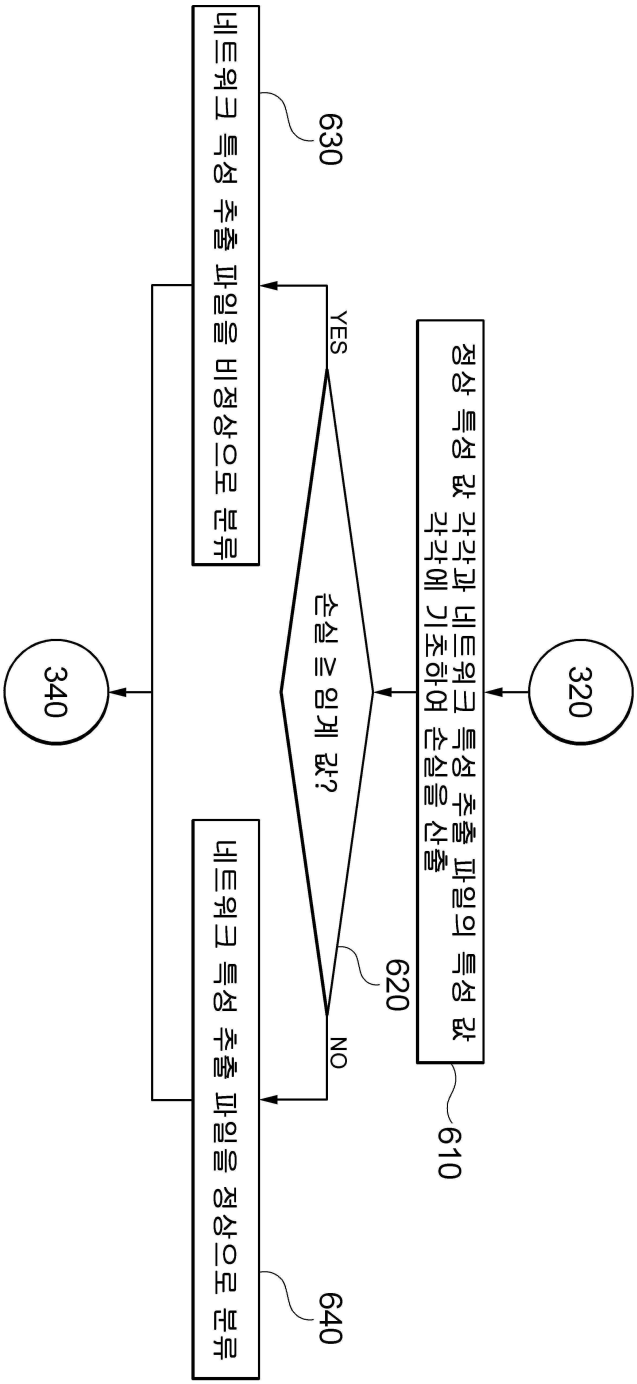
도면4


시 작
네트워크 패킷 캡쳐 파일을 하나 이상의 네트워크 패킷 플로우로 군집화
410
네트워크 특성 추출 파일을 생성
420
학습을 통해 생성된 복수의 네트워크 특성 분류 모델을 저장
430
네트워크 특성 추출 파일에 대한 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델을 결정
440
정상-비정상 분류를 수행
450
네트워크에 대한 공격을 탐지
460
종 료

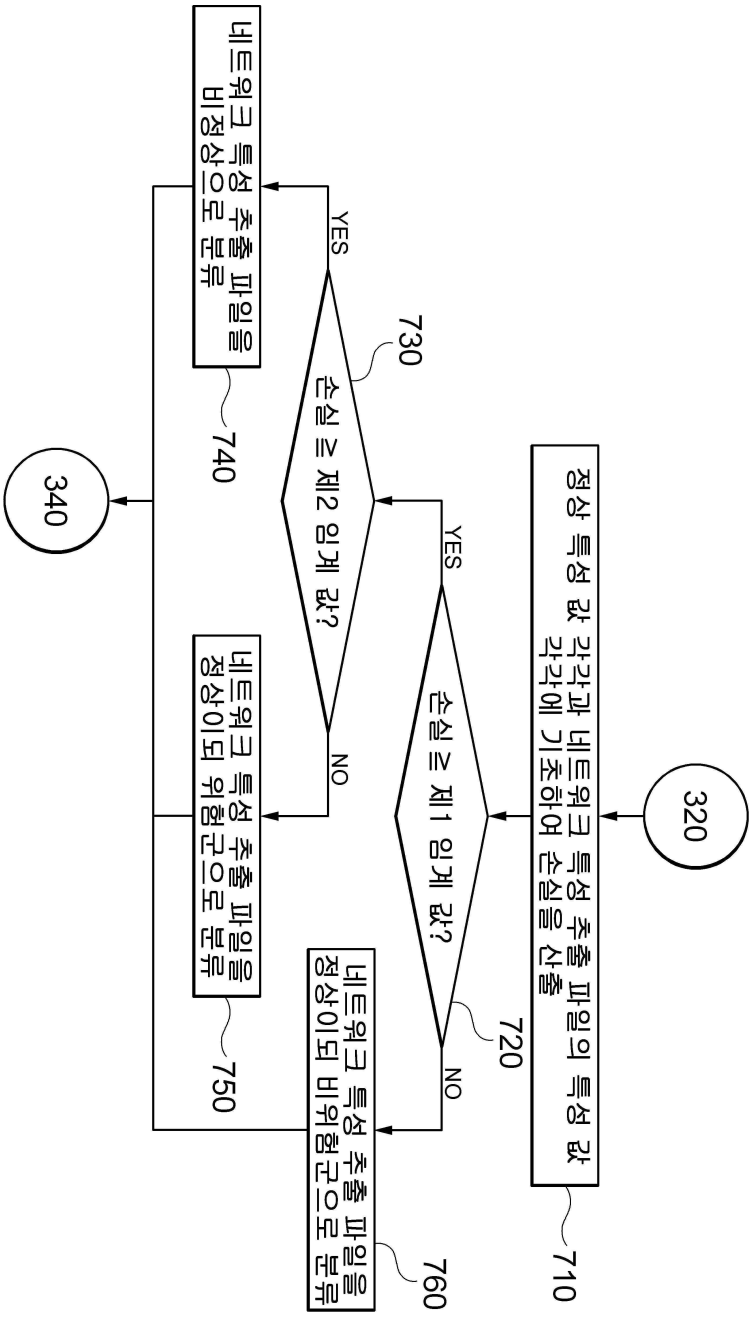
도면5


310
테스트 파일의 분류 결과와 정상-비정상 레이블에 기초하여 모델 각각의 정밀도 및 재현율을 산출
510
정밀도 및 재현율에 기초하여 F1-score를 산출
520
F1-score가 가장 높은 모델을 정상-비정상 분류를 위해 이용할 네트워크 특성 분류 모델로 결정
530
330

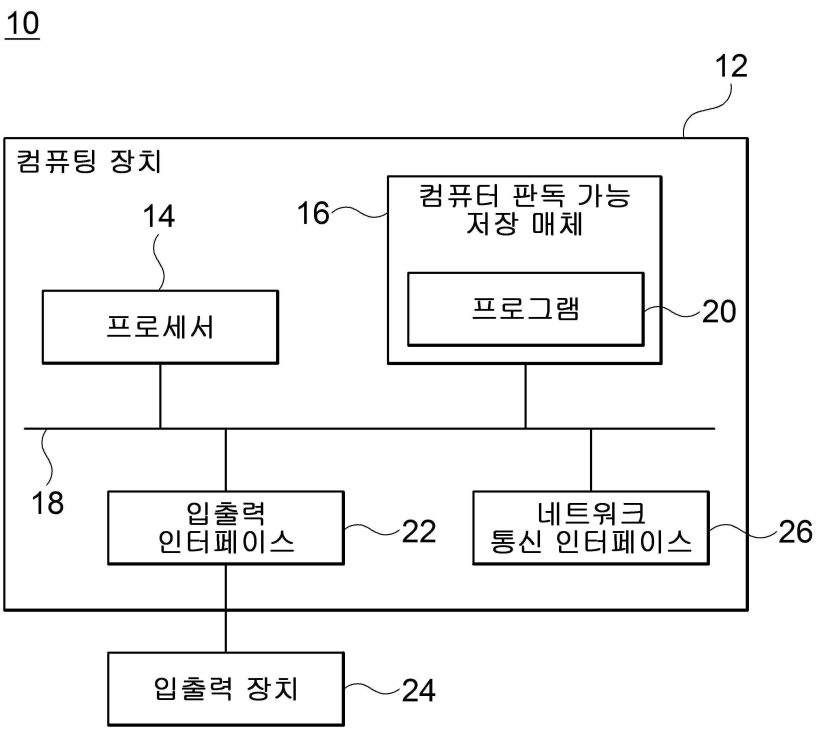
도면6


320
정상 특성 값 각각과 네트워크 특성 추출 파일의 특성 값 각각에 기초하여 손실을 산출
610
손실 ≥ 임계 값?
620
YES
NO
네트워크 특성 추출 파일을 비정상으로 분류
630
네트워크 특성 추출 파일을 정상으로 분류
640
340

도면7


320
정상 특성 값 각각과 네트워크 특성 추출 파일의 특성 값 각각에 기초하여 손실을 산출
710
손실 ≥ 제1 임계 값?
720
YES
NO
네트워크 특성 추출 파일을 정상이되 비위협군으로 분류
760
손실 ≥ 제2 임계 값?
730
YES
NO
네트워크 특성 추출 파일을 비정상으로 분류
740
네트워크 특성 추출 파일을 정상이되 위협군으로 분류
750
340

도면8


10
12
컴퓨팅 장치
14
프로세서
16
컴퓨터 판독 가능
저장 매체
프로그램
20
18
입출력
인터페이스
22
네트워크
통신 인터페이스
26
입출력 장치
24