(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0142443
(43) 공개일자 2021년11월25일

(51) 국제특허분류(Int. Cl.)
G06N 20/00 (2019.01) *H04L 29/06* (2006.01)
(52) CPC특허분류
G06N 20/00 (2021.08)
H04L 63/1416 (2013.01)
(21) 출원번호 10-2020-0059273
(22) 출원일자 2020년05월18일
심사청구일자 2020년05월18일

(71) 출원인
국방과학연구소
대전광역시 유성구 북유성대로488번길 160 (수남동)
(72) 발명자
박정찬
대전광역시 유성구 북유성대로488번길 160
신동일
서울특별시 광진구 능동로 209
(뒷면에 계속)
(74) 대리인
제일특허법인(유), 박장원

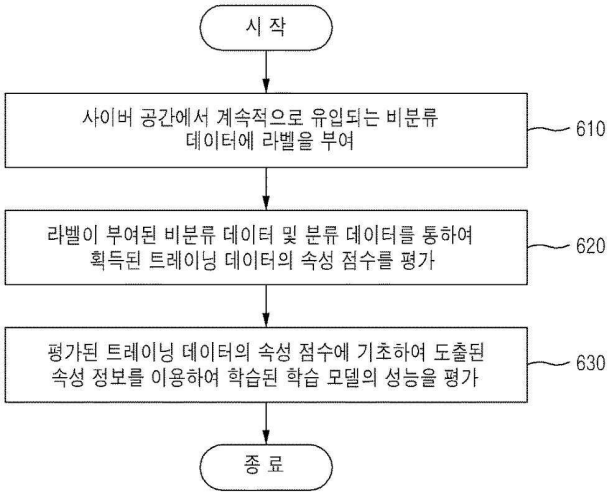
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 **사이버 공간에서 실시간 공격 탐지를 위한 시간에 따른 지속적인 적응형 학습을 제공하는 방법 및 시스템**

(57) 요 약

일 실시예에 따른 탐지 시스템에 의해 수행되는 사이버 공간에서의 실시간 공격을 탐지하는 방법이 제공된다. 상기 방법은 사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여하는 단계; 상기 라벨이 부여된 비분류 데이터 및 분류 데이터를 통하여 획득된 트레이닝 데이터의 속성 점수를 평가하는 단계; 및 상기 평가된 트레이닝 데이터의 속성 점수에 기초하여 도출된 속성 정보를 이용하여 학습된 학습 모델의 성능을 평가하는 단계를 포함할 수 있다.

대 표 도 - 도6



(72) 발명자

신동규

서울특별시 광진구 능동로 209

유지훈

서울특별시 광진구 능동로 209

김진국

대전광역시 유성구 북유성대로488번길 160

명 세 서

청구범위

청구항 1

탐지 시스템에 의해 수행되는 사이버 공간에서의 실시간 공격을 탐지하는 방법에 있어서, 상기 방법은

사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여하는 단계;

상기 라벨이 부여된 비분류 데이터 및 분류 데이터를 통하여 획득된 트레이닝 데이터의 속성 점수를 평가하는 단계; 및

상기 평가된 트레이닝 데이터의 속성 점수에 기초하여 도출된 속성 정보를 이용하여 학습된 학습 모델의 성능을 평가하는 단계를 포함하는 사이버 공간에서의 실시간 공격 탐지 방법.

청구항 2

제1항에 있어서,

상기 사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여하는 단계는,

상기 라벨이 부여된 비분류 데이터 및 분류 데이터에 대한 전처리를 수행하는 단계를 포함하는 사이버 공간에서의 실시간 공격 탐지 방법.

청구항 3

제2항에 있어서,

상기 사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여하는 단계는,

상기 비분류 데이터만으로 파라미터를 생성하는 비지도 프리 트레이닝(Unsupervised pre training)을 수행하고,

상기 분류 데이터를 이용하여 분류기(classifier)를 생성하여 역전파(Backpropagation)를 진행하는 지도 파인 튜닝(supervised fine-tunning)을 수행하여 학습 파라미터를 튜닝하는 것을 특징으로 하는 사이버 공간에서의 실시간 공격 탐지 방법.

청구항 4

제3항에 있어서,

상기 사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여하는 단계는,

자율 학습(self-taught Learning)과 표현 학습(Representation Learning)을 결합한 DCAE(Dilated Convolution Auto encoder) 알고리즘에 기반하여 비분류 데이터로부터 생성된 상기 학습 파라미터를 사용하여 원본 데이터를 변환하고, 상기 변환된 원본 데이터에 대한 재학습을 진행하는 단계를 포함하는 사이버 공간에서의 실시간 공격 탐지 방법.

청구항 5

제2항에 있어서,

상기 사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여하는 단계는,

상기 분류 데이터에 정규화 과정을 적용하고, 상기 비분류 데이터로부터 특징을 추출하고, 상기 추출된 특징에 라벨을 부여함에 따라 생성된 비분류 데이터에 정규화 과정을 적용하는 단계를 포함하는 사이버 공간에서의 실시간 공격 탐지 방법.

청구항 6

제1항에 있어서,

상기 라벨이 부여된 비분류 데이터 및 분류 데이터를 통하여 획득된 트레이닝 데이터의 속성 점수를 평가하는 단계는,

상기 획득된 트레이닝 데이터의 속성 점수를 평가하기 위한 속성별 가중치를 평가하고, 상기 평가된 속성별 가중치에 기초하여 상기 획득된 트레이닝 데이터의 속성 점수를 내림차순으로 정렬하는 단계를 포함하는 사이버 공간에서의 실시간 공격 탐지 방법.

청구항 7

제1항에 있어서,

상기 평가된 트레이닝 데이터의 속성 점수에 기초하여 도출된 속성 정보를 이용하여 학습된 학습 모델의 성능을 평가하는 단계는,

상기 평가된 트레이닝 데이터의 속성 점수에 기초하여 도출된 속성 정보를 이용하여 학습 모델을 학습시키고, 상기 학습된 학습 모델의 성능 평가를 통해 속성 조합을 탐색하는 단계를 포함하고,

상기 속성 조합을 탐색하는 단계에서, 상기 학습된 학습 모델의 성능 평가를 통해 원본 데이터에서 불필요한 속성을 제거하여 상기 속성 조합을 탐색하는 것을 특징으로 하는 사이버 공간에서의 실시간 공격 탐지 방법.

청구항 8

탐지 시스템에 있어서,

사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여하는 자동 라벨링 모듈;

상기 라벨이 부여된 비분류 데이터 및 분류 데이터를 통하여 획득된 트레이닝 데이터의 속성 점수를 평가하는 특징 가중치 연산 모듈; 및

상기 평가된 트레이닝 데이터의 속성 점수에 기초하여 도출된 속성 정보를 이용하여 학습된 학습 모델의 성능을 평가하는 평가 모델 모듈을 포함하는 탐지 시스템.

청구항 9

제8항에 있어서,

상기 자동 라벨링 모듈은,

상기 라벨이 부여된 비분류 데이터 및 분류 데이터에 대한 전처리를 수행하는 것을 특징으로 하는 탐지 시스템.

청구항 10

제9항에 있어서,

상기 자동 라벨링 모듈은,

자율 학습(self-taught Learning)과 표현 학습(Representation Learning)을 결합한 DCAE(Dilated Convolution Auto encoder) 알고리즘에 기반하여 비분류 데이터에로부터 생성된 학습 파라미터를 사용하여 원본 데이터를 변환하고, 상기 변환된 원본 데이터에 대한 재학습을 진행하는 것을 특징으로 하는 탐지 시스템.

청구항 11

제9항에 있어서,

상기 자동 라벨링 모듈은,

상기 분류 데이터에 정규화 과정을 적용하고, 상기 비분류 데이터로부터 특징을 추출하고, 상기 추출된 특징에 라벨을 부여함에 따라 생성된 비분류 데이터에 정규화 과정을 적용하는 것을 특징으로 하는 탐지 시스템.

청구항 12

제8항에 있어서,

상기 특징 가중치 연산 모듈은,

상기 획득된 트레이닝 데이터의 속성 점수를 평가하기 위한 속성별 가중치를 평가하고, 상기 평가된 속성별 가중치에 기초하여 상기 획득된 트레이닝 데이터의 속성 점수를 내림차순으로 정렬하는 것을 특징으로 하는 탐지 시스템.

청구항 13

제8항에 있어서,

상기 평가 모델 모듈은,

상기 평가된 트레이닝 데이터의 속성 점수에 기초하여 도출된 속성 정보를 이용하여 학습 모델을 학습시키고, 상기 학습된 학습 모델의 성능 평가를 통해 속성 조합을 탐색하고,

상기 속성 조합의 탐색은 상기 학습된 학습 모델의 성능 평가를 통해 원본 데이터에서 불필요한 속성을 제거하여 이루어지는 것을 특징으로 하는 탐지 시스템.

발명의 설명

기 술 분 야

[0001] 본 발명은 적응형 학습을 제공하는 방법 및 시스템에 관한 것이다. 보다 상세하게는 사이버 공간에서 실시간 공격 탐지를 위한 시간에 따른 지속적인 적응형 학습을 제공하는 방법 및 시스템에 관한 것이다.

배 경 기 술

[0002] 최근 전 세계적으로 사이버 공격이 개인, 기업, 국가 등을 가리지 않고 다양한 공격이 발생하여 금전적인 피해뿐 아니라 중요한 정보의 탈취 등으로 큰 피해를 보고 있다. 과거 재래식 무기를 이용하는 군은 현재 첨단 기술을 기반으로 고성능 무기를 사용하고 있으며, 첨단 기술의 핵심은 정보통신기술로 각종 컴퓨터, 센서, 네트워크 등을 통해 군의 네트워크를 연결하여 실시간으로 모든 전투 요소를 통합적으로 관리하고 있다. 이에 따라 현대전쟁 영역이 육지, 해상, 공중 이외의 우주, 사이버 공간이 추가된 5개의 영역으로 변경되었다.

[0003] 물리적으로 구분 가능한 공간이 아닌 가상의 사이버 공간은 다른 영역에서의 활동에 대한 자유를 보장하면서 각 영역의 활동을 연결하는 매개체 역할을 한다. 이러한 사이버 공간이 중요한 영역이 되면서 현실 또한 대부분 컴퓨터 시스템에 의존하게 변화시킨다.

[0004] 비특허문헌 Cisco, V. N. I. "Cisco Visual Networking Index: Forecast and Trends, 2017-2022" White paper, 2018>에 따르면 2022년에 도달하게 되면 1인당 네트워크 장치는 3.6개, IP 네트워크에 연결된 장치의 수는 전 세계 인구의 3배가 되며, 전 세계 인터넷 트래픽은 연간 4.8 ZB(Zetta Byte)이상 도달할 것으로 예상하였다. 이와 같이 처리하기 힘들 정도의 많은 양의 사이버 공간의 정보가 실시간으로 발생하면서 수많은 위협에 대하여 지속적인 연구와 발전이 이루어지고 있지만, 보안에 대한 절대적인 안정을 취할 수 없다는 문제점이 있다. 또한, 사이버 공간에서 발생하는 위협은 공공 기관, 민간 기관, 정부와 상업 군대와 비군사적 구분이 모호하기 때문에 이를 분석할 수 있는 운영 시스템을 필요로 한다.

발명의 내용

해결하려는 과제

[0005] 본 발명은 상기와 같은 종래 기술의 기술적 문제를 해결하기 위한 것으로, 본 발명의 목적은 사이버 공간에서 실시간 공격 탐지를 위한 시간에 따른 지속적인 적응형 학습을 제공하는 방법 및 시스템을 제공하는 것이다.

[0006] 또한, 본 발명의 목적은 사이버 공간에서 계속적으로 유입되는 데이터를 분석하기 위하여 특징 선택(Feature Selection)과 연속 학습(Continuous Learning) 알고리즘을 사용하는 CALOT(Continuous Adaptive Learning Over Time)를 운영하는 시스템 및 방법을 제공하는 것이다.

과제의 해결 수단

[0007] 상기와 같은 과제를 해결하기 위한 일 실시예에 따른 탐지 시스템에 의해 수행되는 사이버 공간에서의 실시간 공격을 탐지하는 방법이 제공된다. 상기 방법은 사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여하는 단계; 상기 라벨이 부여된 비분류 데이터 및 분류 데이터를 통하여 획득된 트레이닝 데이터의 속성 점

수를 평가하는 단계; 및 상기 평가된 트레이닝 데이터의 속성 점수에 기초하여 도출된 속성 정보를 이용하여 학습된 학습 모델의 성능을 평가하는 단계를 포함할 수 있다.

[0008] 일 실시 예에 따르면, 상기 사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여하는 단계는, 상기 라벨이 부여된 비분류 데이터 및 분류 데이터에 대한 전처리를 수행하는 단계를 포함할 수 있다.

[0009] 일 실시 예에 따르면, 상기 사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여하는 단계는, 상기 비분류 데이터만으로 파라미터를 생성하는 비지도 프리 트레이닝(Unsupervised pre training)을 수행하고, 상기 분류 데이터를 이용하여 분류기(classifier)를 생성하여 역전파(Backpropagation)를 진행하는 지도 파인 튜닝(supervised fine-tunning)을 수행하여 학습 파라미터를 튜닝할 수 있다.

[0010] 일 실시 예에 따르면, 상기 사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여하는 단계는, 자율 학습(self-taught Learning)과 표현 학습(Representation Learning)을 결합한 DCAE(Dilated Convolution Auto encoder) 알고리즘에 기반하여 비분류 데이터로부터 생성된 상기 학습 파라미터를 사용하여 원본 데이터를 변환하고, 상기 변환된 원본 데이터에 대한 재학습을 진행하는 단계를 포함할 수 있다.

[0011] 일 실시 예에 따르면, 상기 사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여하는 단계는, 상기 분류 데이터에 정규화 과정을 적용하고, 상기 비분류 데이터로부터 특징을 추출하고, 상기 추출된 특징에 라벨을 부여함에 따라 생성된 비분류 데이터에 정규화 과정을 적용하는 단계를 포함할 수 있다.

[0012] 일 실시 예에 따르면, 상기 라벨이 부여된 비분류 데이터 및 분류 데이터를 통하여 획득된 트레이닝 데이터의 속성 점수를 평가하는 단계는, 상기 획득된 트레이닝 데이터의 속성 점수를 평가하기 위한 속성별 가중치를 평가하고, 상기 평가된 속성별 가중치에 기초하여 상기 획득된 트레이닝 데이터의 속성 점수를 내림차순으로 정렬하는 단계를 포함할 수 있다.

[0013] 일 실시 예에 따르면, 상기 평가된 트레이닝 데이터의 속성 점수에 기초하여 도출된 속성 정보를 이용하여 학습된 학습 모델의 성능을 평가하는 단계는, 상기 평가된 트레이닝 데이터의 속성 점수에 기초하여 도출된 속성 정보를 이용하여 학습 모델을 학습시키고, 상기 학습된 학습 모델의 성능 평가를 통해 속성 조합을 탐색하는 단계를 포함할 수 있다. 상기 속성 조합을 탐색하는 단계에서, 상기 학습된 학습 모델의 성능 평가를 통해 원본 데이터에서 불필요한 속성을 제거하여 상기 속성 조합을 탐색

[0014] 다른 실시 예에 따른 탐지 시스템이 개시된다. 상기 탐지 시스템은, 사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여하는 자동 라벨링 모듈; 상기 라벨이 부여된 비분류 데이터 및 분류 데이터를 통하여 획득된 트레이닝 데이터의 속성 점수를 평가하는 특징 가중치 연산 모듈; 및 상기 평가된 트레이닝 데이터의 속성 점수에 기초하여 도출된 속성 정보를 이용하여 학습된 학습 모델의 성능을 평가하는 평가 모델 모듈을 포함할 수 있다.

[0015] 일 실시 예에 따르면, 상기 자동 라벨링 모듈은, 상기 라벨이 부여된 비분류 데이터 및 분류 데이터에 대한 전처리를 수행할 수 있다.

[0016] 일 실시 예에 따르면, 상기 자동 라벨링 모듈은, 자율 학습(self-taught Learning)과 표현 학습(Representation Learning)을 결합한 DCAE(Dilated Convolution Auto encoder) 알고리즘에 기반하여 비분류 데이터에로부터 생성된 학습 파라미터를 사용하여 원본 데이터를 변환하고, 상기 변환된 원본 데이터에 대한 재학습을 진행할 수 있다.

[0017] 일 실시 예에 따르면, 상기 자동 라벨링 모듈은, 상기 분류 데이터에 정규화 과정을 적용하고, 상기 비분류 데이터로부터 특징을 추출하고, 상기 추출된 특징에 라벨을 부여함에 따라 생성된 비분류 데이터에 정규화 과정을 적용할 수 있다.

[0018] 일 실시 예에 따르면, 상기 특징 가중치 연산 모듈은, 상기 획득된 트레이닝 데이터의 속성 점수를 평가하기 위한 속성별 가중치를 평가하고, 상기 평가된 속성별 가중치에 기초하여 상기 획득된 트레이닝 데이터의 속성 점수를 내림차순으로 정렬할 수 있다.

[0019] 일 실시 예에 따르면, 상기 평가 모델 모듈은, 상기 평가된 트레이닝 데이터의 속성 점수에 기초하여 도출된 속성 정보를 이용하여 학습 모델을 학습시키고, 상기 학습된 학습 모델의 성능 평가를 통해 속성 조합을 탐색할 수 있다. 상기 속성 조합의 탐색은 상기 학습된 학습 모델의 성능 평가를 통해 원본 데이터에서 불필요한 속성을 제거하여 이루어질 수 있다.

발명의 효과

[0020] 일 실시 예에 따른 사이버 공간에서 실시간 공격 탐지를 위한 시간에 따른 지속적인 적응형 학습을 제공하는 탐지 방법 및 시스템은 다음과 같은 효과를 가진다.

[0021] 일 실시예에 따른 탐지 시스템은 최적화된 속성을 제시하여, 초기에 들어오는 데이터에서 공격 징후를 분류하는데 최적의 속성들만 남기기 때문에 실시간 공격 징후 탐지를 위해 사용하는 데이터의 양을 감소시키며, 모델을 학습하는 과정에서 비용절감 및 과적합 문제를 해결할 수 있다.

[0022] 일 실시예에 따른 탐지 시스템은 연속 학습(Continuous Learning)을 통해 학습된 대표 패턴을 기억하고 새로 들어오는 패턴에만 학습하여, 매번 전체를 학습하지 않고 추가된 데이터만 학습하여, 과거 및 현재 데이터에 대해 분석할 수 있는 정교한 모델을 완성시킬 수 있다.

[0023] 상술한 본 발명의 특징 및 효과는 첨부된 도면과 관련한 다음의 상세한 설명을 통하여 보다 분명해 질 것이며, 그에 따라 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다.

도면의 간단한 설명

[0024] 도 1은 일 실시예에 따른 탐지 시스템에서 전처리 과정을 설명하기 위한 도면이다.

도 2는 일 실시예에 따른 탐지 시스템에서 사이버 공간에서의 실시간 공격을 탐지하기 위한 구조를 설명하기 위한 도면이다.

도 3은 일 실시예에 따른 탐지 시스템에서 자동 라벨링 동작을 설명하기 위한 도면이다.

도 4는 일 실시예에 따른 탐지 시스템에서 특징 가중치 연산 모듈의 구조를 설명하기 위한 도면이다.

도 5는 일 실시예에 따른 탐지 시스템에서 평가 모델 모듈의 구조를 설명하기 위한 도면이다.

도 6은 일 실시예에 따른 탐지 시스템에서 사이버 공간에서의 실시간 공격을 탐지하는 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0025] 상술한 본 발명의 특징 및 효과는 첨부된 도면과 관련한 다음의 상세한 설명을 통하여 보다 분명해 질 것이며, 그에 따라 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다. 본 발명은 다양한 변경을 가할 수 있고 여러 가지 형태를 가질 수 있는바, 특정 실시 예들을 도면에 예시하고 본문에 상세하게 설명하고자 한다. 그러나 이는 본 발명을 특정한 개시형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 본 명세서에서 사용한 용어는 단지 특정한 실시 예들을 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다.

[0026] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시 예를 가질 수 있는바, 특정 실시 예들을 도면에 예시하고 상세한 설명에 구체적으로 설명하고자 한다. 그러나 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0027] 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용한다.

[0028] 제1, 제2등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.

[0029] 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. "및/또는" 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.

[0030] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미가 있다.

[0031] 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않아야 한다.

[0032] 이하의 설명에서 사용되는 구성요소에 대한 접미사 "모듈", "블록" 및 "부"는 명세서 작성의 용이함만이 고려되어 부여되거나 혼용되는 것으로서, 그 자체로 서로 구별되는 의미 또는 역할을 갖는 것은 아니다.

[0033] 이하, 본 발명의 바람직한 실시 예를 첨부한 도면을 참조하여 당해 분야에 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 설명한다. 하기에서 본 발명의 실시 예를 설명함에 있어, 관련된 공지의 기능 또는 공지의 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략한다.

[0034] 이하에서는, 본 명세서에 따른 무기체계 환경/신뢰성시험용 이종 시험장비 인터페이스 방법 및 장치에 대해 설명하기로 한다.

[0035] 본 명세서에서는 사이버 공간에서 계속해서 유입되는 데이터를 분석하기 위해서 기존의 일반적인 기계 학습이 아닌 특징 선택(Feature Selection)과 연속 학습(Continuous Learning) 알고리즘을 사용하는 CALOT(Continuous Adaptive Learning Over Time) 운영 시스템 및 방법에 대하여 설명하기로 한다. 실시간으로 이루어지는 사이버 공간에서 지속적이며, 반복적으로 운영되는 모델에서 이전 데이터를 학습된 모델에 새로운 데이터가 들어오면 기존 학습 모델에 이어서 학습하기 위하여 연속 학습이 가능한 점진적 학습(Incremental Learning)을 이용할 수 있다. 점진적 학습은 이미 학습한 대표 패턴을 기억하고 새로 들어오는 패턴에 대해서만 학습하는 방법으로, 추가로 학습 패턴이 발생할 경우 매번 전체를 학습하지 않고 새로 추가된 데이터만 학습한다. 이는 시간과 자원이 제한된 상황에도 학습이 가능하며, 풍부한 자원이 제공되는 경우 배치(Batch)단위로 분할하여 학습을 하기 때문에 실시간으로 이루어지는 사이버 공간 침입 탐지에 적합한 학습 방법이다.

[0036] 이와 관련하여, 도 1은 일 실시예에 따른 탐지 시스템에서 전처리 과정을 설명하기 위한 도면이다.

[0037] 탐지 시스템은 CALOT 테스트에 사용하는 데이터 세트(Dataset)에 대한 전처리 과정을 수행할 수 있다. 데이터 세트로 CSE-CIC-IDS2018를 사용할 수 있다. CSE-CIC-IDS 2018 데이터 세트와 관련하여, 비특허문헌 2<Sharafaldin, Iman, et al. "Towards a reliable intrusion detection benchmark dataset." Software Networking 2018.1 (2018): 177-200.>를 참조할 수 있다. 일 예로, CSE-CIC-IDS 2018 데이터 세트는 이러한 intrusion detection benchmark dataset에 의하여 만들어진 데이터 세트와 연관될 수 있다. 비특허문헌 2에 따르면, 기존에 많은 사람들이 사용하였던 DARPA98, KDD99, ISC2012 및 ADFA13과 같은 기존의 데이터 세트의 신뢰도에 대해서 검증하였다, 하지만 실제로 트래픽 부족, 볼륨 부족, 공격의 다양성 불충분, 현재 추세를 반영할 수 없는 패킷 정보, 페이로드 익명화, 기능 집합과 메타 데이터 부족과 같은 이유로 모든 테스트한 데이터 세트가 신뢰할 수 없다는 결과가 도출될 수 있다. 이에 CSE(Communications Security Establishment)와 캐나다 사이버 보안 연구소 공동 프로젝트로 위의 다양한 기준을 만족하는 7개의 일반적인 공격 네트워크 흐름을 포함하는 신뢰할 수 있는 데이터 집합을 생성하였다. 데이터 세트는 하루 단위로 구성되어 있으며, 컴퓨터 당 네트워크 트래픽(Pcap)과 이벤트 로그를 포함한 원시 데이터를 기록하였다. 원시 데이터의 특징(Feature)을 추출하기 위해서 종래에는 NetMate라는 도구를 사용하여 특징을 생성하였지만, CSE-CIC-IDS2018 데이터 세트에서는 자체적으로 제작한 오픈 소스 툴인 CICFlowMeter를 사용하여 복수 개(예를 들면, 80개) 이상의 트래픽 특징을 추출할 수 있다. 결과적으로 포괄적인 네트워크 트래픽 기능 세트 및 기계 학습 알고리즘의 성능을 평가하여 특정 공격 범주를 탐지하기 위한 최상의 기능 집합을 나타내는 데이터 세트가 완성되었다. 일 실시예에 따른 탐지 시스템은 CSE-CIC-IDS2018 데이터 세트를 사용할 수 있다.

[0038] 탐지 시스템은 CALOT를 통해 데이터를 학습하기 위하여 전처리를 수행할 수 있다. 이때, 80% 이상 시간을 데이터 수집 및 전처리에 사용하기 때문에, 매우 중요한 과정이다. 도 1을 참고하면, 전처리 과정에 대한 전체적인 구조가 도시되어 있다.

[0039] 일 실시예에 따르면, 전처리 과정에서 필터 접근(Filter Approach)과 래퍼 접근(Wrapper Approach)를 혼합한 하이브리드 특징 접근(Hybrid Feature Approach)을 사용할 수 있다.

[0040] 필터 접근이란 일반적으로 전처리 과정에서 사용되며, 머신 러닝 알고리즘과 독립적으로 적용될 수 있다. 이는 필터 접근의 결과가 머신 러닝에 있어서 최선의 특징 서브셋(Best Feature Subset)이 아니라, 결과 변수와 상관관계에 근거하여 다양한 통계 점수를 도출하여 속성에 대해 랭킹 순위를 부여한다는 뜻이다. 이를 통해 머신 러닝 사용자는 학습 이전에 특징 랭크(Feature-Rank)를 통해 각각의 특징이 얼만큼의 영향력을 가지는지에 대한

정보를 가지고 특징 서브셋(Feature Subset)을 구성할 수 있다.

[0041] 래퍼 접근은 특징 서브셋을 이용하여 모델을 학습시켜 나가며, 추론을 기반으로 이전 모델에서 특징을 가감할 것인지 여부를 결정한다. 다만, 추론 기반으로 모델을 비교해 나가다 보니 계산적으로 많은 비용이 소모될 수 있다. 필터 접근이 머신 러닝의 전처리 과정에서 모델의 성능과 독립적으로 특징 랭크를 추출하는 것에 비해, 특징 서브셋을 이용하여 머신 러닝을 진행하면서 높은 성능을 보이는 특징을 가감하기 때문에 머신 러닝 알고리즘과 직접적인 연관이 있어 성능 측면에서 훨씬 더 좋은 효과를 보인다.

[0042] 탐지 시스템은 전처리 과정을 데이터에 따라 2가지로 구분할 수 있다. 첫 번째로, 분류 데이터(Labeled Data)(101)로 CSE-CIC-IDS2018 데이터 세트이다. CSE-CIC-IDS2018 데이터 세트는 이미 가공되어 사용자를 통해 라벨링(labeling)된 데이터로 이미 가공된 데이터이기 때문에 기계 학습을 바로 적용할 수 있는 Data Construction을 가진 CSV 파일 형식을 가진다. 탐지 시스템은 분류 데이터에 다른 사전 과정이 없이 정규화(Normalization) 과정(130)을 적용할 수 있다.

[0043] 본 명세서에서 개시되는 데이터는 분류 데이터(101) 이외에 비분류 데이터(Unlabeled Data)를 포함할 수 있다. 이와 관련하여, 비분류 데이터(Unlabeled Data)는 Raw Network Packet(pcap)을 통해 입력되는 데이터(111)일 수 있다. Raw Network Packet(pcap)을 통해 입력되는 데이터(111)를 Raw Network Packet 데이터(111)로 지칭할 수 있다. 일반적으로 사이버 공간에서 수집되는 데이터들은 라벨링(Labeling)이 되어 있지 않은 비분류 데이터 형태이다. 이는 사이버 공간뿐만 아니라 다른 분야에서도 동일하며, 가공되지 않은 데이터는 일반적으로 라벨링이 되어 있지 않다. 이에, 비분류 데이터에 대한 라벨(Label)을 부여해줄 필요가 있다.

[0044] 탐지 시스템은 사이버 공간에서 수집되는 Raw Network Data에서 특징을 추출하기 위해서 CICFlowMeter(112)를 이용할 수 있다. 이때, CICFlowMeter(112)를 이용함에 따라 추출된 데이터는 라벨이 없이 특징으로 구성된 비분류 데이터(Unlabeled Data)(113)가 된다. 탐지 시스템은 비분류 데이터(113)에 CALOT 자동 라벨링 모듈(Automatic Labeling Module)을 이용한 자동 라벨링(114)을 통하여 라벨을 부여할 수 있다. 자동 라벨링 모듈에 대한 설명은 도 2에서 상세하게 설명하기로 한다. 자동 라벨링을 통해 성공적으로 라벨이 부여된 데이터는 분류 데이터와 같이 표 1에 도시된 데이터 구성(Data Construction)(120)을 가지게 된다. 데이터 구성은 라벨을 포함한 복수 개(예를 들면, 79개)의 특징을 가지고 있는 데이터 구조로, 학습 가능한 형태이지만 정규화(Normalization)가 되어있지 않기 때문에 학습할 경우 많은 시간 소요와 부정확한 결과값이 나올 가능성이 크다.

표 1

[0045]

No	Destination Port	Flow Duration	Total Forward Packets	...	Idle mean	Idle std	Label
1	54865	3	11	...	0	0	BENIGN
2	55054	109	3	...	0	0	BENIGN
3	55055	52	1	...	542	607	BENIGN
4	46236	11509095	2	...	36300000	42800000	DoS
5	54863	3	2	...	0	0	BENIGN
6	54925	2	1	...	0	0	FTPPatator
7	54208	10279821	3	...	36100000	43100000	DoS
8	80	5006127	4	...	0	0	Injection
9	54925	1022	6	...	0	0	BENIGN
10	54925	1022	7	...	2	4	BENIGN
11	32148	48	8	...	3	9	BENIGN
12	46236	651	1	...	18	35	BENIGN
13	80	6642489	1	...	5871254	6578416	BruteForce
14	35842	142	8	...	556	48	Bot
15	18736	26	2	...	0	0	BENIGN
	...	...	...	...	...	...	...

[0046] 탐지 시스템은 보다 정확한 학습 결과와 빠른 학습을 위하여 분류 데이터 및 비분류 데이터 각각에 대한 정규화(Normalization)(130)를 진행할 수 있다. 정규화는 4가지 단계로 구성될 수 있다. 정규화에서 Null Value Delete는 데이터 구성(Data Construction)에서 각각의 특징으로부터 널 벨류(Null value)를 제거하는 것으로, 널 벨류를를 통해 부정확한 학습이 되는 것을 방지할 수 있다. Incorrect Value Delete는 비정상적인 값을 제거하는 것으로, '인피니티', 'NaN'과 같은 값을 제거 혹은 대체한다. 대체해주는 경우 여러 가지 방법이 존재하는데 일반적으로 많이 사용하는 방법은 '0채우기', '가장 많은 빈도수 데이터', '가장 높은값', '가장

낮은값', '평균값'으로 대체한다. Min-max Normalization는 데이터들의 분포의 차이가 기 설정된 기준 이상(예를 들면, 큰)일 경우, 학습에 있어 많은 부하가 발생하기 때문에 데이터의 값을 일정 범위의 분포로 스켈링(scaling)하는 것으로, 일반적으로 0 내지 1 사이 값을 많이 사용하며 수학식 1은 다음과 같다.

수학식 1

[0047] $$\text{MinMax} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

[0048] 특징 선택(Feature Selection)은 전체 특징에서 학습에 가장 좋은 성능의 특징 서브셋(Feature Subset)을 생성하는 것이다. 특징 선택이란 머신 러닝(Machine learning)을 통해 풀고자 하는 문제에 대한 사전 지식이 있다면 올바른 정보만 모을 수 있지만, 대부분의 경우 사전 지식이 없기에 불필요한 정보가 포함되는 경우가 많다. 이에, 특징 선택은 사전 지식 없이 수집된 모든 속성에 대한 부분 집합을 이용하여 효율적인 속성 집합을 생성하는 방법을 의미한다. 다시 말해서, 다양한 속성 중 결과에 영향이 없다 판단되는 속성을 제거하는 방법으로 원본 데이터에서 불필요한 속성을 제거하여, 최적의 모델을 생성해 나가는 과정이다.

[0049] 특징 선택은 특징 가중치 연산 모듈(Feature Weight Operation module)을 통해 진행되며, 이에 대한 자세한 설명은 도 2에서 설명하기로 한다. 탐지 시스템은 정규화의 4가지 단계를 수행함에 따라 생성된 최종의 학습 데이터(트레이닝 데이터)(140) 예시는 표 2와 같이 나타낼 수 있다. 이와 관련하여, 표 2는 정규화 데이터 구조(Normalization Data Structure)를 나타낸다.

표 2

[0050]

No	Destination Port	Flow Duration	Total Forward Packets	...	Idle mean	Idle std	Label
1	0.00427	0.00..033	0.00..184	...	0	0	BENIGN
2	0.00478	0.00..474	0.00..033	...	0	0	BENIGN
3	0.00479	0.00..296	0.00..016	...	0.00..841	0.00..957	BENIGN
4	0.00401	1.0	0.00..025	...	0.315	0.371	DoS
5	0.00425	0.00..033	0.00..025	...	0	0	BENIGN
6	0.00447	0.00..025	0.00..016	...	0	0	FTPPatator
7	0.00439	0.893	0.00..033	...	0.313	0.374	DoS
8	0.00..062	0.434	0.00..048	...	0	0	Injection
9	0.00447	0.00..102	0.00..062	...	0	0	BENIGN
10	0.00447	0.00..102	0.00..077	...	0.00..025	0.00..048	BENIGN
11	0.00279	0.00..117	0.00..089	...	0.00..033	.000..091	BENIGN
12	0.00401	0.00..992	0.00..016	...	0.00..217	.0.00..422	BENIGN
13	0.00..062	0.0577	0.00..016	...	0.511	0.571	BruteForce
14	0.00311	0.00..693	0.00..089	...	0.00..866	0.00..589	Bot
15	0.00162	0.00..377	0.00..025	...	0	0	BENIGN
	...	...	...	...	...	...	...

[0051] 도 2는 일 실시예에 따른 탐지 시스템에서 사이버 공간에서의 실시간 공격을 탐지하기 위한 구조를 설명하기 위한 도면이다. 한편, 도 3은 일 실시예에 따른 탐지 시스템에서 자동 라벨링 동작을 설명하기 위한 도면이다. 도 4는 일 실시예에 따른 탐지 시스템에서 특징 가중치 연산 모듈의 구조를 설명하기 위한 도면이다. 또한, 도 5는 일 실시예에 따른 탐지 시스템에서 평가 모델 모듈의 구조를 설명하기 위한 도면이다.

[0052] 도 2를 참조하면, 탐지 시스템은 자동 라벨링 모듈(Automatic Labeling Module)(210), 특징 가중치 연산 모듈(220), 평가 모델 모듈(230)을 포함할 수 있다.

[0053] 탐지 시스템은 자동 라벨링 모듈(Automatic Labeling Module)(210), 특징 가중치 연산 모듈(220), 평가 모델 모듈(230)을 이용하여 사이버 공간에서 계속적으로 유입되는 데이터를 분석할 수 있다.

[0054] 자동 라벨링 모듈(210)은 비분류 데이터(UnLabeled Data)에 대한 라벨을 자동으로 부여할 수 있다. 자동 라벨링 모듈(210)은 탐지 시스템의 가장 처음 시작인 부분에서 진행될 수 있다. 자동 라벨링 모듈(210)에서 자동으로 라벨을 부여하기 위해 자율 학습(self-taught Learning)과 표현 학습(Representation Learning)의 개념을 결합

한 DCAE(Dilated Convolution Auto encoder) 알고리즘을 사용할 수 있다. DCAE는 자율 학습과 같이 비분류 데이터에서 생성된 학습 파라미터를 사용하여 원본 데이터를 변환하고, 변환된 원본 데이터에 다시 학습을 진행하는 비지도 학습(Unsupervised Learning)의 한 형태이다. 라벨이 없이 특징만을 사용하여 파라미터를 학습하는 특징 때문에 비지도 특징 학습(Unsupervised Feature Learning)이라고도 부른다. 도 3을 참고하면, 비분류 데이터만으로 파라미터를 생성하는 비지도 프리 트레이닝(Unsupervised pre training)(310)을 수행한 후, 소수의 분류 데이터를 이용하여 분류기(classifier)를 생성하여 역전파(Backpropagation)를 진행하는 지도 파인 튜닝(supervised fine-tunning)(320)을 수행하여 기존 파라미터를 튜닝(parameter tuning)함으로써 더 정교한 파라미터를 완성 시키는 구조를 가진다.

[0055] 도 3에 도시된 바와 같이, 스택 확장형 컨볼루션 자동 인코더 구조(Stack Dilated Convolution Auto Encoder Structure)를 나타낸 것이다. DCAE에서 오토인코더(Autoencoder)를 구성할 때 일반적인 컨볼루션(Convolution)대신 확장형 컨볼루션(Dilated Convolution)을 사용한다. 그 이유로 레이어(Layer)를 풀링(Pooling)하지 않는 구조를 통해 기존의 컨볼루션보다 적은 파라미터를 사용하여 빠른 학습 속도를 보여주며, 필터 내부에 제로 패딩(Zero Padding)을 추가하여 강제로 리셉티브 필드(Receptive Field)를 늘리기 때문에 공간 차원(Spatial Dimension)의 손실이 적고 대부분 가중치가 0이기 때문에 연산 효율이 좋다.

[0056] 특징 가중치 연산 모듈(220)은 각 속성별 점수를 평가하고, 모든 데이터를 사용한 모델에 대한 초기 성능을 측정하도록 구성될 수 있다. 즉, 특징 가중치 연산 모듈(220)을 통해 각 속성별 점수를 평가하고, 모든 데이터를 사용한 모델에 대한 초기 성능을 측정하는 방법이 수행될 수 있다. 도 4를 참고하면, 특징 가중치 연산 모듈의 구조를 설명하기 위한 도면에 관한 것으로, 특징 가중치 연산(Feature Weight Operation)(410) 및 특징 평가 풀(Pool of Feature Evaluation)(420)이 도시되어 있다. 특징 가중치 연산(410)을 통해 특징 선택 알고리즘(Feature Selection Algorithm), 교차 검증(cross validation), 특징 평가(feature evaluation)를 수행할 수 있다. 또한, 특징 가중치 연산(410)을 통해 정확도 예측(Estimate Accuracy)을 수행할 수 있다. 또한, 특징 가중치 연산(410)을 통해 높은 정확도 특징(High Accuracy Feature)을 선택하고, 최적 특징(Optimal Feature)을 선택할 수 있다.특징 평가 풀(420)과 관련하여, 거리(distance), 정보(information), 의존성(dependency), 일관성(consistency) 및 특징 중요도(feature importance) 중 적어도 하나 이상을 고려하여 선택된 특징을 평가할 수 있다.

[0057] 한편, 특징 가중치 연산 모듈(220)은 각 속성별 점수를 평가하고, 모든 데이터를 사용한 모델에 대한 초기 성능을 측정함에 있어, 도 4의 특징 가중치 연산(410)과 특징 평가 풀(420)과 연관된 기능 또는 동작이 수행될 수 있다. 하지만, 특징 가중치 연산 모듈(220)은 도 4에서 제시되는 특징 가중치 연산(410)과 특징 평가 풀(420)과 연관된 기능 또는 동작에 한정되지 않고, 다음과 같이 다양한 형태의 특징 선택을 이용하여 데이터의 각 속성 별 점수를 평가할 수 있다.

[0058] 특징 가중치 연산 모듈(220)은 특징 선택을 이용하여 데이터의 각 속성 별 점수를 평가할 수 있다. 특징 가중치 연산 모듈(220)은 각 속성별 가중치를 평가하게 되며, 평가된 각 속성별 가중치에 기초하여 각 속성별 점수를 내림차순으로 정렬한다. 특징 가중치 연산 모듈(220)은 속성(List of Features) 또는/및 속성 집합(List of Feature sets)에 따라 각 속성에 대한 점수를 기반으로 내림차순으로 정렬할 것인지 또는 속성 집합에 따른 점수를 기반으로 내림차순으로 정렬할 것인지 선택할 수 있다. 각 속성에 대한 점수를 기반으로 정렬하는 경우, 각각의 속성별로 평가된 점수를 기반으로 정렬하게 된다.

[0059] 또한, 속성 집합에 따른 점수를 기반으로 정렬하는 경우, 순열(permutation)된 특징 셋(Feature set)별로 점수를 부여하는 방법을 적용할 수 있다. 이때, 사용되는 값은 속성의 개수(Length, n)와 평가 점수(Value, v)를 이용하며, 각 값을 정규화를 통해 0 ~ 1의 값으로 만들어준다. 수학식2처럼 속성의 개수가 적게 쓰일수록 높게 평가하고, 집합의 점수가 높을수록 높게 평가한다. 추가적으로, 이후 평가 모델 모듈(230)에서 성능을 비교하기 위해, 모든 속성을 이용하여, 학습한 모델의 성능 평가가 진행될 수 있다. 이때, 10-fold cross validation방법을 이용하여 성능을 도출해낼 수 있다.

수학식 2

[0060] $$W_n = \frac{\left(1-\frac{L_n-L_{min}}{L_{max}-L_{min}}\right)+\left(\frac{V_n-V_{min}}{V_{max}-V_{min}}\right)}{2}$$

[0061] 평가 모델 모듈(230)은 특징 가중치 연산부(220)를 통하여 도출된 속성을 이용하여, 실제 학습 모델에 적용하여 학습 이후 성능 평가를 통해 최적의 속성 조합을 찾을 수 있다. 도 5를 참고하면, 평가 모델(510) 구조를 설명하기 위한 도면이다. 도 5의 평가 모델(510)은 도 2의 평가 모델 모듈(230)에 대응할 수 있다. 도 2의 평가 모델 모듈(230)은 도 5의 평가 모델(510)과 연동하여 성능 평가를 수행할 수 있다. 도 2 및 도 5를 참조하면, 평가 모델 모듈(230)은 임계값을 예측(Estimate Threshold), 최적의 임계값(Best Threshold)을 선택하고, 이에 따른 평가 모델을 선택(Select Model)할 수 있다.

[0062] 도 5를 참조하면, 평가 모델 모듈(230)은 수학식 3을 이용하여 임계값(Threshold Value)(520)을 도출할 수 있다. 도출된 임계값을 통해, 시간, 메모리와 성능이 적절한 학습 모델을 에이전트가 스스로 선택할 수 있다. 이후 10-fold cross validation을 통해 조합에 대한 성능지표를 도출하게 된다.

수학식 3

[0063] $$Threshold\,Value = \frac{M+R+P+L}{4}$$

[0064] 결과적으로 특징 가중치 연산 모듈(220)과 평가 모델 모듈(230)를 통해 모델에 최적화된 속성을 제시하여, 초기에 들어오는 데이터에서 공격 징후를 분류하는데 최적의 속성들만 남기기 때문에 실시간 공격 징후 탐지를 위해 사용하는 데이터의 양이 줄어들며, 모델을 학습하는 과정에서 비용절감 및 과적합 문제를 해결할 수 있다. 연속 학습(Continuous Learning)을 통해 학습된 대표 패턴을 기억하고 새로 들어오는 패턴에만 학습하여, 매번 전체를 학습하지 않고 추가된 데이터만 학습하여, 과거 및 현재 데이터에 대해 분석할 수 있는 정교한 모델이 완성될 수 있다.

[0065] 도 6은 일 실시예에 따른 탐지 시스템에서 사이버 공간에서의 실시간 공격을 탐지하는 방법을 설명하기 위한 흐름도이다.

[0066] 단계(610)에서 탐지 시스템은 사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여할 수 있다. 탐지 시스템은 라벨이 부여된 비분류 데이터 및 분류 데이터에 대한 전처리를 수행할 수 있다. 탐지 시스템은 자율 학습(self-taught Learning)과 표현 학습(Representation Learning)을 결합한 DCAE(Dilated Convolution Auto encoder) 알고리즘에 기반하여 비분류 데이터로부터 생성된 학습 파라미터를 사용하여 원본 데이터를 변환하고, 변환된 원본 데이터에 대한 재학습을 진행할 수 있다. 탐지 시스템은 분류 데이터에 정규화 과정을 적용하고, 비분류 데이터로부터 특징을 추출하고, 추출된 특징에 라벨을 부여함에 따라 생성된 비분류 데이터에 정규화 과정을 적용할 수 있다.

[0067] 단계(620)에서 탐지 시스템은 라벨이 부여된 비분류 데이터 및 분류 데이터를 통하여 획득된 트레이닝 데이터의 속성 점수를 평가할 수 있다. 탐지 시스템은 획득된 트레이닝 데이터의 속성 점수를 평가하기 위한 속성별 가중치를 평가하고, 평가된 속성별 가중치에 기초하여 획득된 트레이닝 데이터의 속성 점수를 내림차순으로 정렬할 수 있다.

[0068] 단계(630)에서 탐지 시스템은 평가된 트레이닝 데이터의 속성 점수에 기초하여 도출된 속성 정보를 이용하여 학습된 학습 모델의 성능을 평가할 수 있다. 탐지 시스템은 평가된 트레이닝 데이터의 속성 점수에 기초하여 도출된 속성 정보를 이용하여 학습 모델을 학습시키고, 학습된 학습 모델의 성능 평가를 통해 속성 조합을 탐색할 수 있다.

[0069] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치

는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0070] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0071] 실시예에 따른 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 실시예를 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.

[0072] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0073] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

[0074] 이상에서는 일 실시 예에 따른 무기체계 환경/신뢰성시험용 이종 시험장비 인터페이스 방법 및 장치에 대해 설명하였다. 일 실시 예에 따른 무기체계 환경/신뢰성시험용 이종 시험장비 인터페이스 방법 및 장치는 다음과 같은 효과를 가진다.

[0075] 일 실시 예에 따른 사이버 공간에서 실시간 공격 탐지를 위한 시간에 따른 지속적인 적응형 학습을 제공하는 탐지 방법 및 시스템은 다음과 같은 효과를 가진다.

[0076] 일 실시예에 따른 탐지 시스템은 최적화된 속성을 제시하여, 초기에 들어오는 데이터에서 공격 징후를 분류하는데 최적의 속성들만 남기기 때문에 실시간 공격 징후 탐지를 위해 사용하는 데이터의 양을 감소시키며, 모델을 학습하는 과정에서 비용절감 및 과적합 문제를 해결할 수 있다.

[0077] 일 실시예에 따른 탐지 시스템은 연속 학습(Continuous Learning)을 통해 학습된 대표 패턴을 기억하고 새로 들어오는 패턴에만 학습하여, 매번 전체를 학습하지 않고 추가된 데이터만 학습하여, 과거 및 현재 데이터에 대해 분석할 수 있는 정교한 모델을 완성시킬 수 있다.

[0078] 상술한 본 발명의 특징 및 효과는 첨부된 도면과 관련한 다음의 상세한 설명을 통하여 보다 분명해 질 것이며, 그에 따라 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다.

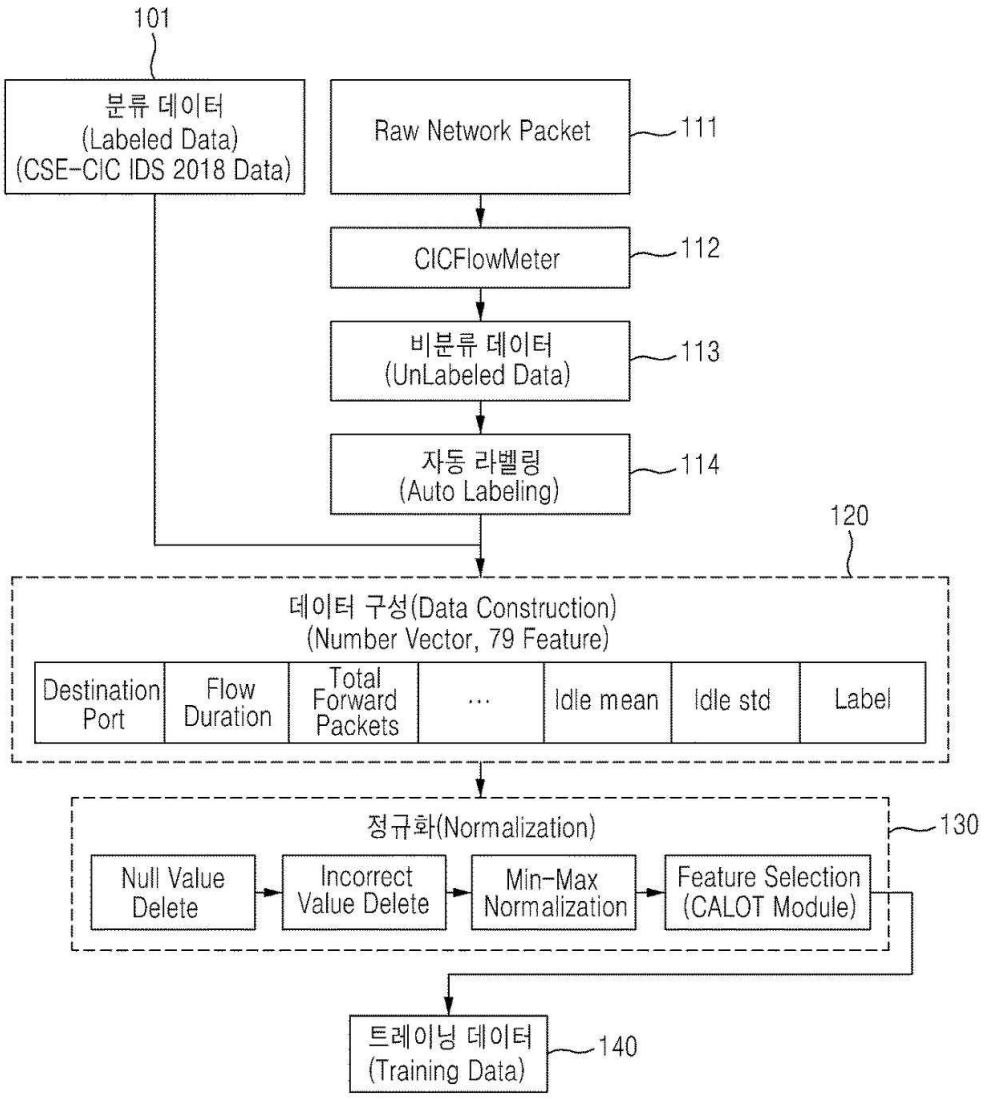
[0079] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시 예를 가질 수 있는바, 특정 실시 예들을 도면에 예시하고 상세한 설명에 구체적으로 설명하고자 한다. 그러나 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0080] 소프트웨어적인 구현에 의하면, 본 명세서에서 설명되는 절차 및 기능뿐만 아니라 각각의 구성 요소들에 대한 설계 및 파라미터 최적화는 별도의 소프트웨어 모듈로도 구현될 수 있다. 적절한 프로그램 언어로 쓰여진 소프

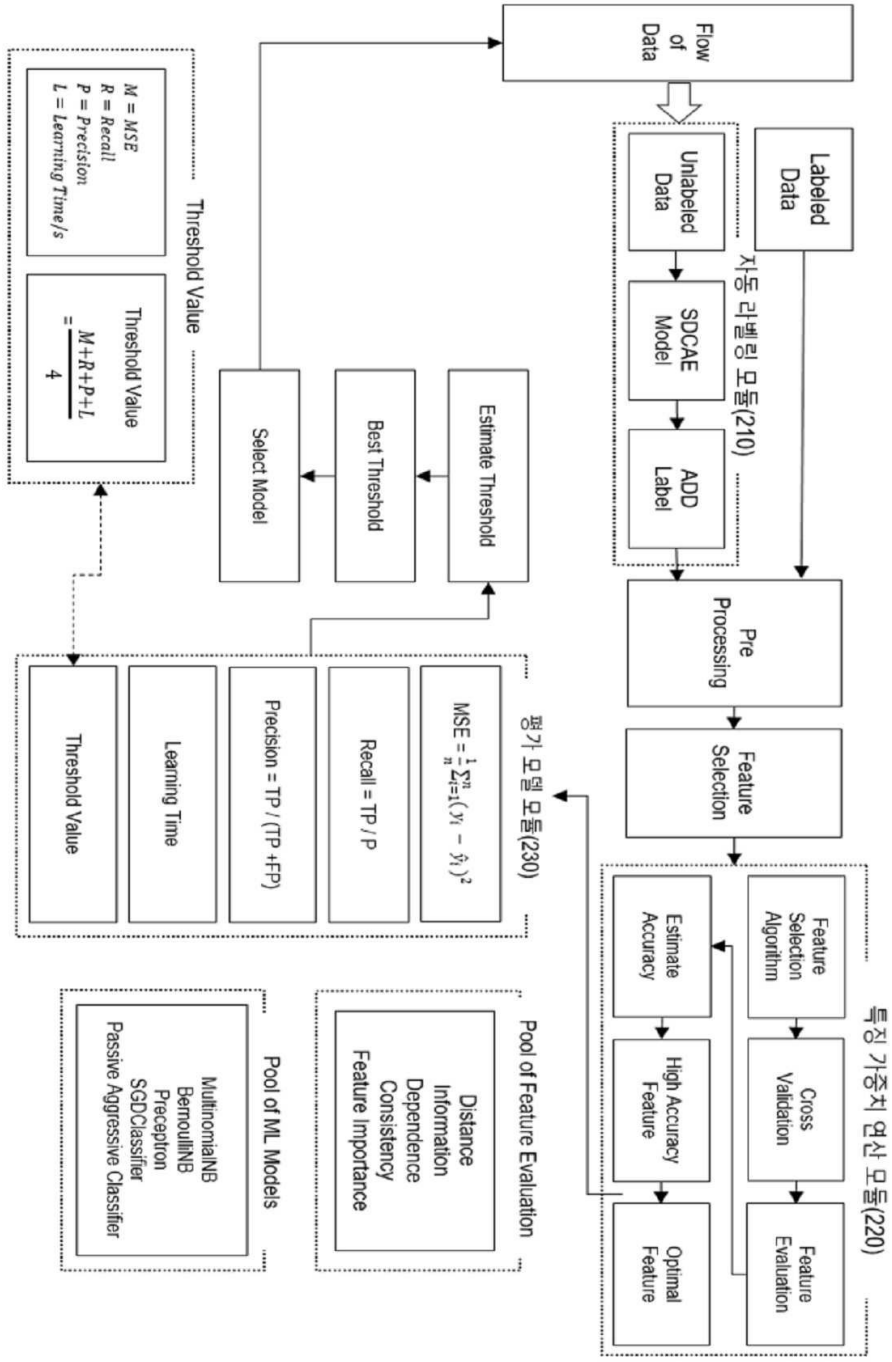
트웨어 어플리케이션으로 소프트웨어 코드가 구현될 수 있다. 상기 소프트웨어 코드는 메모리에 저장되고, 제어부(controller) 또는 프로세서(processor)에 의해 실행될 수 있다.

도면

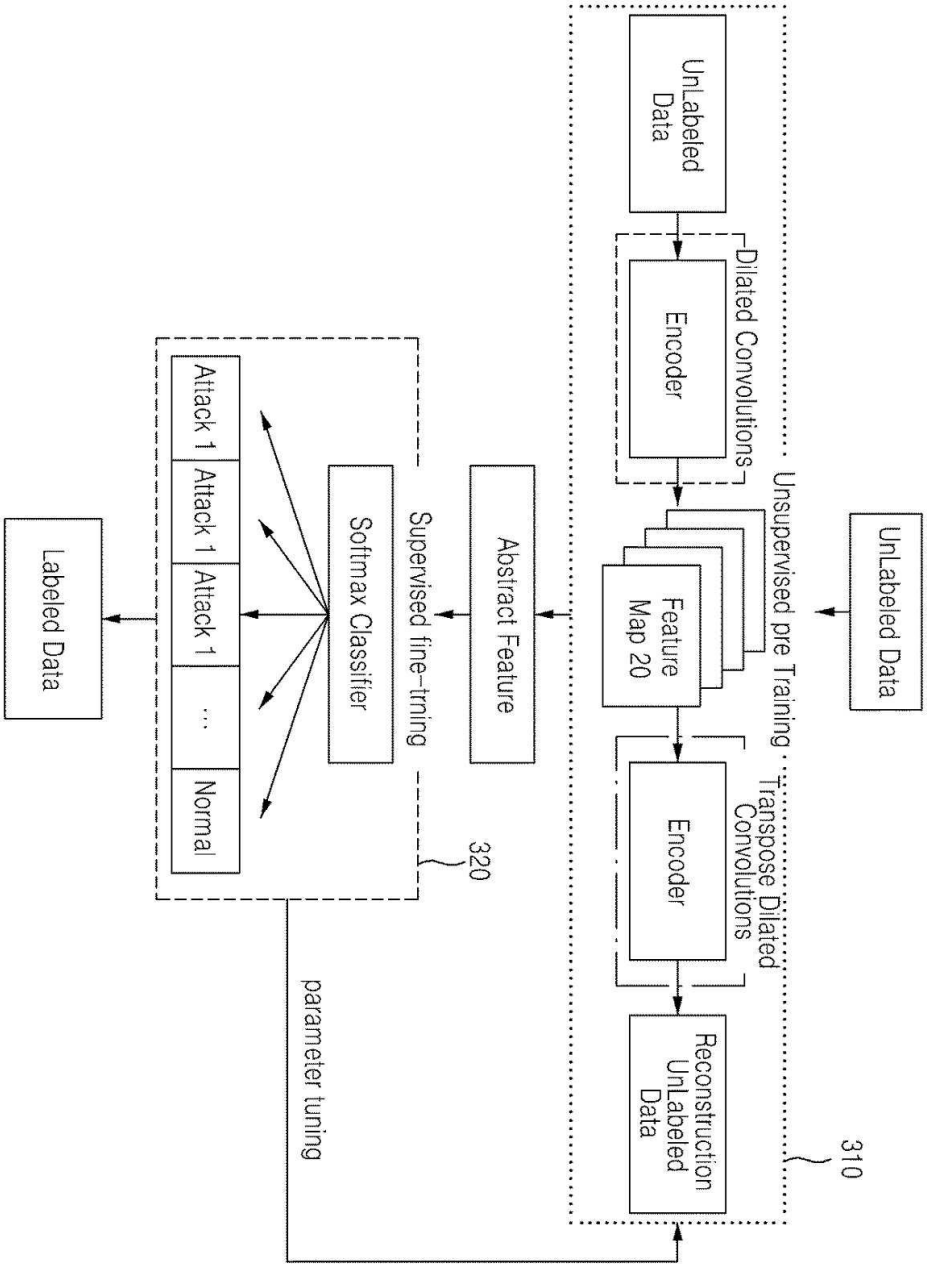
도면1



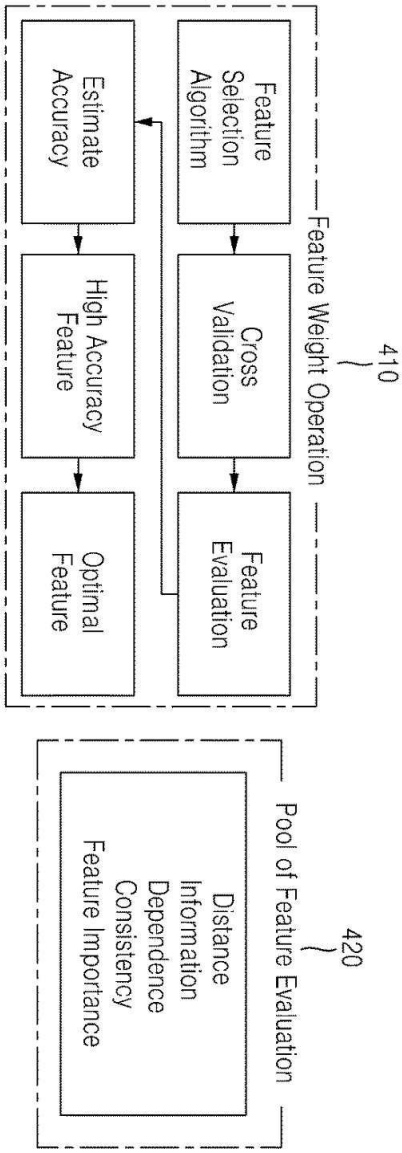
도면2


Flow of Data
Labeled Data
Unlabeled Data
SDCAE Model
ADD Label
자동 라벨링 모듈(210)
Pre Processing
Feature Selection
Feature Selection Algorithm
Estimate Accuracy
Cross Validation
High Accuracy Feature
Feature Evaluation
Optimal Feature
특징 가중치 연산 모듈(220)
Estimate Threshold
Best Threshold
Select Model
Threshold Value
Learning Time
Precision = TP / (TP +FP)
Recall = TP / P
MSE = $\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$
평가 모델 모듈(230)
M = MSE
R = Recall
P = Precision
L = Learning Time/s
Threshold Value
Threshold Value = $\frac{M+R+P+L}{4}$
Distance
Information
Dependence
Consistency
Feature Importance
Pool of Feature Evaluation
MultinomialNB
BernoulliNB
Preceptron
SGDClassifier
Passive Aggressive Classifier
Pool of ML Models

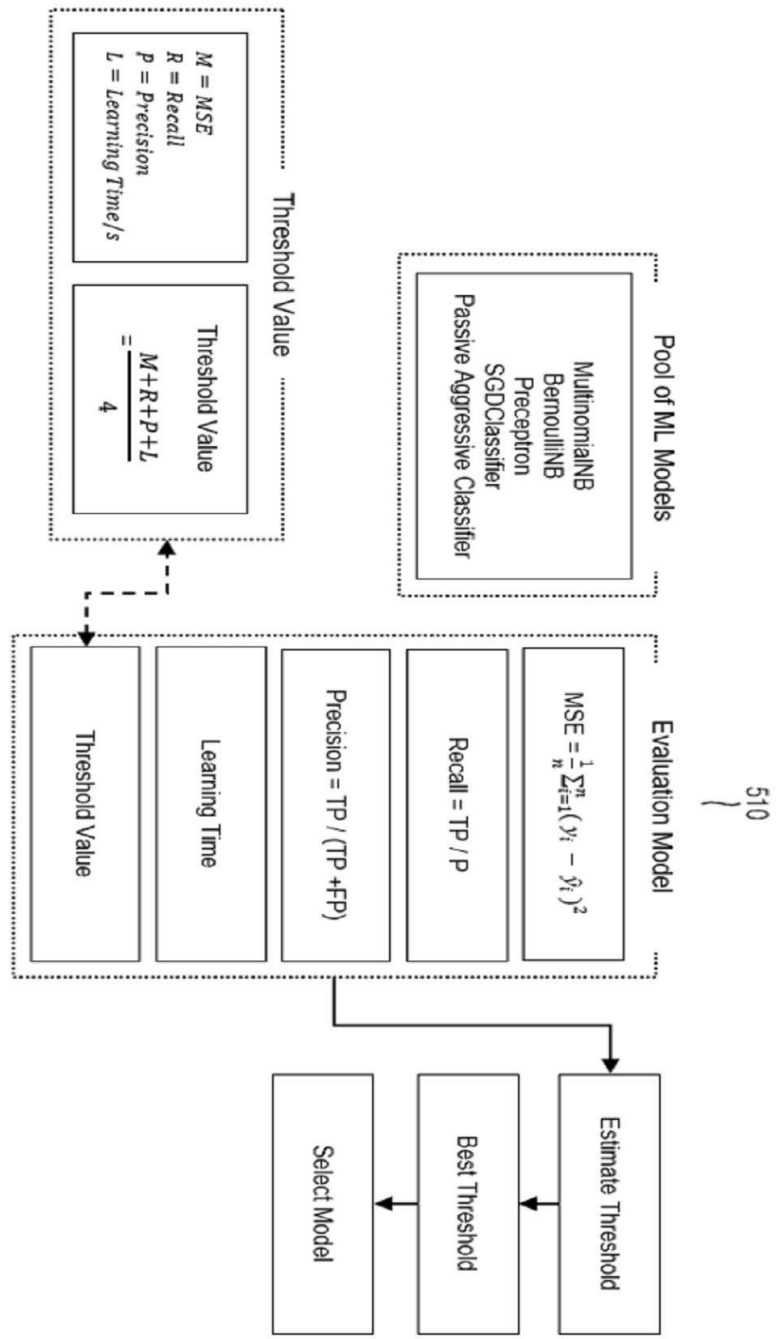
도면3


UnLabeled Data
Dilated Convolutions
Encoder
Feature Map 20
Unsupervised pre Training
UnLabeled Data
Transpose Dilated Convolutions
Encoder
Reconstruction UnLabeled Data
310
Abstract Feature
Softmax Classifier
Supervised fine-trning
Attack 1
Attack 1
Attack 1
...
Normal
320
Labeled Data
parameter tuning

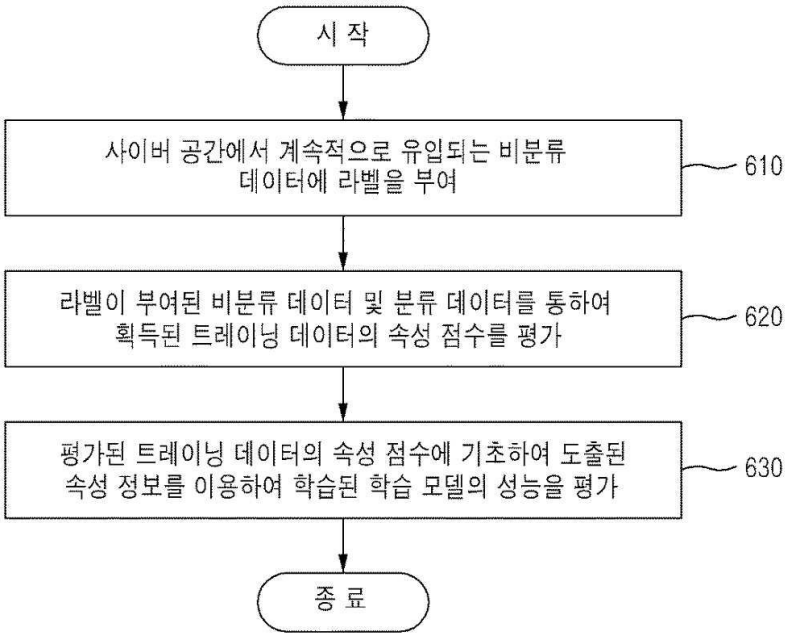
도면4


Feature Weight Operation
410
Feature Selection Algorithm
Cross Validation
Feature Evaluation
Estimate Accuracy
High Accuracy Feature
Optimal Feature
Pool of Feature Evaluation
420
Distance
Information
Dependence
Consistency
Feature Importance

도면5


510
Pool of ML Models
MultinomialNB
BernoulliNB
Preceptron
SGDClassifier
Passive Aggressive Classifier
Evaluation Model
MSE = $\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$
Recall = TP / P
Precision = TP / (TP +FP)
Learning Time
Threshold Value
Threshold Value
M = MSE
R = Recall
P = Precision
L = Learning Time/s
Threshold Value
$= \frac{M+R+P+L}{4}$
Estimate Threshold
Best Threshold
Select Model

도면6


시 작
사이버 공간에서 계속적으로 유입되는 비분류 데이터에 라벨을 부여
610
라벨이 부여된 비분류 데이터 및 분류 데이터를 통하여 획득된 트레이닝 데이터의 속성 점수를 평가
620
평가된 트레이닝 데이터의 속성 점수에 기초하여 도출된 속성 정보를 이용하여 학습된 학습 모델의 성능을 평가
630
종 료