딥뉴럴네트워크 기반의 네트워크 침입탐지시스템 설계

A Design of Deep Neural network-based Network Intrusion Detection System

권현, 방승호, 박기웅[1)]
Hyun Kwon, Seungho Bang, Ki-Woong Park

(01805) 서울특별시 노원구 화랑로 574 육군사관학교 전자공학과
(04353) 서울시 용산구 이태원로 22 합동참모본부 전력기획부
(05006) 서울특별시 광진구 능동로 209 세종대학교 정보보호학과
hkwon1@kma.ac.kr, bsh9468@gmail.com, woongbak@sejong.ac.kr

요 약

최근 들어, 딥뉴럴네트워크는 이미지 인식, 패턴 분석, 침입탐지 등 다양한 분야에 활용되고 있다. 본 논문에서는 딥뉴럴네트워크 기반의 침입탐지방법에 대하여 여러 가지 파라미터를 조절하여 탐지 성공률을 실험적으로 분석하였다. 대표적인 침입 데이터인 KDD CUP 99 데이터셋을 사용하였으며, 텐서플로우 머신러닝 라이브러리를 이용하였고 각 파라미터를 수정하여 탐지율을 측정하였다. 추가적으로 선형회귀(linear regression, LR), 나이브 베이즈 분류 (naive bayes classification, NB), 최근접 이웃(k-nearest neighbors, KNN), 의사결정트리(decision tree, DT), 랜덤 포레스트(random forest, RF) 등의 방법들에 대해서도 실험적으로 비교 분석하였다. 각 파라미터로는 각 층의 노드 수, Dropout의 수, 활성화 함수 등을 고려하여 분석하였으며, 은닉층이 5개인 딥뉴럴네트워크일 때, 93.8% 탐지 정확도로 가장 높은 것을 확인할 수 있었다.

Abstract

Recently, deep neural networks (DNNs) provide the good performance for image recognition, pattern analysis and intrusion detection. In this paper, we analyzed experimentally the success rate of detection by manipulating various parameters for intrusion detection method of deep neural network. We used the KDD CUP 99 data set, which is a widely used intrusion data, and modified each parameter using the TensorFlow machine learning library to measure the detection accuracy. We also analyzed linear regression (LR), naive bayes classification (NB), k-nearest neighbors (KNN), decision trees (DT) and random forests. (random forest, RF). We analyzed the parameters by modifying the number of nodes in each layer, the number of dropouts, and the activation function. Experimental results show that DNNS has the highest detection with 93.8% accuracy with 5 hidden layers.

1) 교신저자

논문투고일자 : 2019. 09. 23 심사일자 : 2019. 09. 24 게재확정일자 : 2020. 02. 21

키워드: 침입탐지시스템, 딥뉴럴네트워크, 기계학습, KDD CUP 99
Keyword: Intrusion detection system, Deep neural network, Machine learning, KDD CUP 99

1. 서론

컴퓨터 기술의 발전을 통해서 정보 통신 기술(information and communication technology, ICT)이 급속도로 발전되고 있다. 하지만 이러한 정보 통신 기술과 더불어 다양하고 복잡한 침입공격들도 다양하게 나타나고 있다. 따라서 이러한 침입공격들로부터 시스템을 안전하게 보호하고 방어하는 침입탐지시스템은 중요하다. 하지만 기존의 침입탐지시스템(intrusion detection system, IDS)[1]은 잘 알려진 파라미터를 기반으로 악의적인 공격자의 행동을 탐지하기 때문에 효과적이지만 새로운 침입에 대하여 상대적으로 취약한 한계점이 있다.

침입탐지시스템은 크게 오용탐지(signature detection)[2]와 이상탐지(anomaly detection)[3]로 구분이 된다. 오용탐지는 사전에 알려진 공격 패턴이나 이상한 징후가 시그너처(signature)를 통해서 침입을 탐지하는 방법을 의미한다. 이 방법은 알려진 공격에 대해서 효과적으로 탐지가 가능하지만 알려지지 않은 공격에 대해서는 침입을 허용하는 한계점이 있다. 반면에 이상탐지는 정상적인 패턴 행동을 분석하고 이상한 변화가 적정수준을 초과하였을 때, 이상징후로 침입을 탐지하는 방법이다. 이 방법은 정상적인 경우에도 오탐지를 할 가능성이 높으며 지속적으로 시스템의 점검이 요구된다. 이처럼 오용탐지와 이상탐지를 이용하여 침입탐지 시스템을 적용하고 있지만 여러 가지 제한사항이 존재한다. 특히, 알려지지 않은 공격에 대해서 이상징후 탐지를 하는 데 있어서 정상적인 상태에 대한 정의를 명확히 하기가 어렵고 지속적으로 새로운 행동 패턴을 네트워크에서 학습하기가 쉽지 않다. 따라서 많은 침입탐지시스템이 높은 잘못탐지(false detection)을 하거나 알려지지 않은 공격에 대해서 탐지율이 낮은 경향이 있다.

최근 딥러닝 알고리즘 중 딥뉴럴네트워크[4]를 이용하여 이미지 인식[5], 음성 인식[6], 패턴 분석[7] 등에 좋은 성능을 보여주고 있다. 딥러닝 알고리즘은 계층적인 구조로 여러 개의 층들이 정보를 처리하는 모델로 구성이 된다. 이 모델은 사람이 직접 파라미터를 설정하는 기존 머신러닝 방법과 다르게 많은 양의 학습데이터를 통하여 모델이 스스로 파라미터를 최적화하는 단계를 갖는다.

이러한 딥러닝 모델을 침입탐지시스템을 적용하면 모델이 새로운 유형 패턴을 스스로 학습하여 최적의 파라미터를 설정함으로써, 침입을 효과적으로 탐지한다. 이 논문에서는 딥뉴럴네트워크를 이용한 침입탐지시스템을 각 파라미터를 조절하여 탐지율을 분석하였다. 이 논문의 공헌점은 다음과 같다. 먼저, 딥뉴럴네트워크를 적용한 침입탐지시스템에서 각 파라미터가 미치는 영향에 대해서 분석하였다. 딥뉴럴네트워크 모델에서 대상이 되는 파라미터는 각 층의 노드수, Dropout, 활성화 함수, epoch 수 등에 의하여 침입탐지시스템의 성능을 분석하였다. 두 번째로 네트워크에서 자주 사용되는 KDD CUP 99 데이터셋[8]을 이용하여 탐지 성능을 측정하였다. 세 번째로는 다른 머신러닝 기법인 선형회귀(linear regression, LR)[9], 나이브 베이즈 분류(naive bayes classification, NB)[10], 최근접 이웃(k-nearest neighbors, KNN)[11], 의사결정트리(decision tree, DT)[12], 랜덤 포레스트(random forest, RF)[13] 등과 비교하여 성능 분석을 하였다.

이 장의 구성은 다음과 같다. 2장에서는 침입탐지시스템에 대한 관련연구를 소개하고 3장에서는 문제정의를 한다. 4장에서는 딥뉴럴네트워크에 대한 구조에 대하여 설명하고 5장에서는 실험 및 분석결과를 보여준다. 6장에서는 논문에 대한 토론을 하고 7장에서는 논문에 대한 결론으로 구성되어 있다.

2. 관련연구

이 장은 딥뉴럴네트워크를 적용한 침입탐지시스템을 이해하기 위하여 딥러닝 모델과 침입탐지시스템에 대한 관련연구를 소개한다. 2.1장에서 침입탐지시스템에 대한 관련연구를 설명하고 2.2장에서 딥러닝 모델에 대한 관련연구를 설명한다.

2.1 침입탐지시스템에 관한 연구

침입 공격으로부터 방어하기 위한 여러 가지 방어 솔루션들[1] [15] 중에 침입탐지시스템[1]은 공격 패턴에 대한 매칭을 이용하여 위협을 탐지하고 차단하는 시스템이다. 이 시스템은 룰(rule) 기반으로 침입을 탐지하기 때문에 상대적으로 잘못탐지(false detection)가 높은 편이다. 이전 침입탐지 연구들에서는 공격 패턴을 다양한 머신러닝 기술을 이용해서 공격 패턴을 정의하고 잘못탐지(false detection)를 줄여나갔다. 이러한 머신 러닝기술을 적용한 방법으로 서포트백터 머신(support vector machine, SVM) [16], 의사결정트리(decision tree, DT) [12], 베이시안 분류(bayesian classification) [10]가 있다. 또한, 악의적인 트래픽 탐지를 위해 k-means 방법을 적용하여 탐지하는 방법도 있다. Shin et al. 연구진[17]은 k-means 알고리즘을 적용하여 DDoS attack과 Witty worm attack을 탐지할 수 있는 파라미터를 설정하고 비계층적인 클러스터링을 통하여 데이터 유사성을 찾는 방법을 제안하였다. Hatim et al. 연구진[18]은 k-means에 SVM 방법을 혼합한 하이브리드 머신러닝 기법을 적용하여 공격을 탐지하는 시스템을 제안하였다. 하지만 이러한 기존 탐지 기법은 과거 패턴 추출과 학습을 분리된 머신러닝 방법을 이용하지만, 개선된 딥뉴럴네트워크를 이용한 방법이 새롭게 소개되고 있다. 이는 이미 알고 있는 룰 기반이거나 악의적인 공격 패턴을 분석하여 탐지하는 방법과 달리 이상징후에 위협과 관련된 대량의 데이터를 통해서 모델 스스로가 직접적인 관계성을 찾는다. Ni et al. 연구진[19]은 deep belief network(DBNs)을 이용하여 침입탐지하는 방법을 보여주었고, SVM 모델보다 좀 더 6% 개선된 성능을 가져왔다. 또 다른 연구로 Jo et al. 연구진[20]은 forward additive neural network (FANN)와 기존 SVM모델의 결과를 비교하여 침입탐지하는 연구를 소개하였다. FANN은 역전파(back-propagation)의 취약점을 보완하여 생성한 알고리즘이다. FANN은 SVM보다 좀 더 좋은 정확도와 탐지율을 보여준다.

딥러닝 모델을 이용한 침입탐지 연구에서 대표적으로 4가지 연구가 있다. 먼저, Kim et al.[21] 연구진에 의해서 딥러닝 모델에 의한 침입탐지시스템에 대하여 연구하였다. 이 연구에서는 supervised, unsupervised, semi-supervised, weakly supervised, reinforcement 등에 대한 다양한 딥러닝모델을 이용한 침입탐지시스템을 분석하였다. 두 번째로, Mathai, K et al.[22] 연구진은 state preserving extreme learning machine(SPELM) 알고리즘을 이용하여 기존 Deep Belief Network (DBN) 알고리즘보다 좀 더 성능이 개선된 것을 보여주었다. SPELM 알고리즘은 얼굴인식, 보행자 인식, 네트워크 침입탐지 인식에 사용될 수 있다. 이 연구진은 NSL-KDD 데이터셋에 대하여 SPELM 방법을 이용해서 기존 DBN이 52.85 성능에 탐지율에 비해 93.2% 탐지율 향상을 가져왔다. 세 번째로, Aggarwal, Preeti et al [23] 연구진은 KDD CUP 99 데이터를 이용하여 침입탐지방법에 대하여 연구하였다. 이 연구에서는 KDD CUP 99 데이터를 분석하여 여러 개의 class label을 분석하였다. 이러한 class label에 대하여 1개부터 4개 등 attribute class을 동시에 고려하여 Random forest, OneR, Naive Bayes 방법에 대하여 탐지율(detection rate)과 잘못탐지(false detection)를 분석하였다. 이러한 분석을 통해서 어떤 attribute가 가장 중요한 가중치가 있는 지 분석을 하였다. 네 번째로, Gurung, S et al [24] 연구진은 NSL-KDD 데이터셋을 이용하여 네트워크 침입탐지에 대해서 연구하였다. 이 연구에서는 sparse auto-encoders 방식으로 패턴에 대해서 학습을 하고

sparse auto-encoders를 통해서 학습된 사용자의 행동패턴을 logistic regression을 통해서 탐지하는 방법을 제안하였다. 이 방법을 통해서 87.2% 탐지율의 성능을 보여주었다.

데이터셋 측면에서, 딥러닝 기술을 이용한 침입탐지 분석연구는 대표적으로 2가지가 있다. 먼저, Özgür, Atilla et al [25] 연구진은 리뷰 논문으로써, 침입탐지에서 사용되는 데이터셋인 KDD CUP 99, NSL KDD, DARPA에 대하여 분석을 하였다. 또한 이러한 데이터를 이용하여 탐지를 적용한 SVM, LR, DT 등의 논문이 몇편의 논문이 쓰여졌는지 분석하였다. 두 번째로, Meena, G et al [26] 연구진은 KDD CUP 99 데이터셋과 NSL-KDD 데이터셋에 대하여 기존 의사결정트리와 비교하여 개선된 의사결정트리를 제안하였다. 하지만 이 방법은 의사결정트리에만 국한되어 있어 있는 한계점이 있다.

유형 측면에 있어서, 딥뉴럴네트워크 방법을 적용한 침입탐지시스템은 시그니처 기반에 탐지기반이 된다. 기존 데이터셋을 통해서 침입과 관련된 패턴분석을 딥뉴럴네트워크가 학습을 하여, 비슷한 유형에 패턴이 나타날 경우 탐지하는 방법이다. 기존 방법과 다른 점은 이러한 패턴 부분을 사람이 직접 모델의 파라미터 등을 설정을 하였지만 딥뉴럴네트워크 방법은 훈련용 데이터를 통해서 모델 스스로가 최적의 파라미터를 선정한다는 차이점이 있다.

2.2 딥러닝 모델에 관한 연구

딥러닝 모델은 인간의 뇌와 구조에서 영감을 받아 만들어낸 수학적 인공지능 모델이다. 다양한 수준의 추상적인 단계 때문에, 입력값의 각 특징들은 결과값과 매칭되어 모델이 학습을 한다. 사람이 조작하는 방법과 다르게 학습하는 과정에서 모델이 학습데이터를 통해서 모델 내부상의 각 최적의 파라미터를 설정한다. 이 연구[27]에서는 1943년 McCulloch와 Pitts에 의해서 처음으로 제안되어 뉴럴네트워크에 대한 개념이 처음으로 등장하였다. 그 이후로 뉴럴네트워크에 대한 역전파 알고리즘과 ReLU(rectified linear unit) 함수[28]를 통해서 더 발전하게 되었다.

이론적으로 딥러닝 모델의 한 층은 다양한 노드와 가중치로 구성되어 있다. 이는 인간의 뉴런과 시냅스의 과정을 모방하여 구성된다. 각 노드는 특정 수준에서 반응하고 이러한 반응은 각 노드의 가중치와 곱 관계식을 통해서 구성되어 있다. 각 노드는 다양한 입력값의 가중치를 가지고 있기 때문에 서로 다른 가중치는 다양한 입력값 안에서 조정되어진다. 따라서 이렇게 계산된 모든 수치들은 활성화 함수를 통해서 결과값으로 분류하거나 regression 분석에 적용된다. 군중행동탐지 [29], 데이터 예측 및 관리 [30], 자율주행차량 [31] 등에 활용되는 딥러닝 모델은 많은 분류, 인식, 예측, 생성 분야에 활용되고 있다.

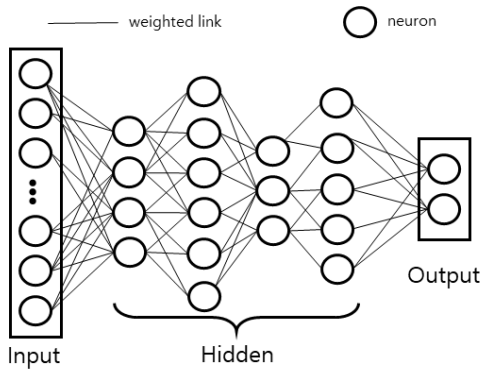
3. 문제정의

딥러닝 모델 중에 딥뉴럴네트워크를 이용한 방법은 다중 퍼셉트론을 이용하여 구성된 모델이다. Vigneswaran et al. 연구진[32]에서는 각 층(layer)에 따라서 침입탐지시스템의 성능을 분석한 내용을 발표하였다. 하지만 딥뉴럴네트워크에 안에는 각 층 뿐만 아니라 각 층의 노드수, 활성화 함수, dropout, epoch 등에 다양한 파라미터가 존재한다. 시스템 관리자 입장에서 어떠한 파라미터가 침입감내시스템에 영향을 주는 지에 대한 내용을 이해할 필요가 있다. 이 연구에서는 여러 가지 파라미터가 침입탐지시스템의 성능에 어떤 영향을 주는 지 분석을 하였고, 다른 머신러닝 기법의 성능에 대해서도 비교하였다.

4. 딥뉴럴네트워크를 이용한 침입탐지시스템

(그림 1)과 같이 딥뉴럴네트워크에 대한 시스템 구성은 입력층(input layer), 은닉층(hidden layer), 출력층(output layer)으로 구성되어 있다. 입력층에서 보면 각 입력값들은 노드에 1:1로 매칭이 된

다. 은닉층(hidden layer)에서는 각 층의 수는 이 모델의 복잡성을 나타낸다. 출력층의 노드의 수는 예측하는 결과값의 유형들을 나타낸다. 은닉층에 있는 노드와 가중치의 결합을 통해서 출력층의 노드에 영향을 준다. 이전 층의 입력 값과 가중치의 곱의 합으로 계산되어 나타난다. 추가적으로 활성화 함수는 이전 층에 노드과 가중치의 합에 의해서 특정 임계치를 넘게 되면 특정값을 주고 그렇지 않으면 0값을 준다. 뉴럴네트워크는 이러한 과정을 gradient descent와 back-propagation을 계산을 통해서 최적의 손실함수를 줄임으로써, 각 층에 있는 최적의 파라미터를 여러 번의 학습과정을 통해서 설정해나간다. 뉴럴네트워크의 성능은 학습데이터에 의존적이기 때문에 학습데이터의 전처리하는 과정이 중요하다.



(그림 1) 딥뉴럴네트워크의 구성

딥뉴럴네트워크에서 구성되는 파라미터와 구성요소들은 활성화 함수(activation function), 드랍아웃(Dropout), 배치 사이즈(batch size), 반복횟수(epoch), 각 층의 노드수(Number of node), 은닉층의 층개수(Number of hidden layers), 최적화 알고리즘 등이 있다. 먼저, 활성화 함수는 step function, sigmoid function[33], ReLU function이 대표적으로 있다. 이 중 sigmoid function과 ReLU를 보편적으로 사용되며, sigmoid function은 0과 1사이의 값만 같는 비선형 함수로써, $S(t)=\frac{1}{1+e^{-t}}$의 함수식을 갖는다. 반면에 ReLU는 gradient vanishing 문제를 해결 하기 위해서 사용되는 것으로 특정 임계치 이상일 때에는 $y=x$의 1차 선형으로 증가하여 gradient vanishing 문제를 해결하는 방법으로 사용된다. 두 번째로, 드랍아웃(Dropout)은 오버피팅(over-fittig)을 막기 위한 방법으로써, 딥뉴럴네트워크가 학습 중일 때, 랜덤하게 특정 노드가 학습하는 것을 방해함으로써, 학습이 특정 데이터에 치중되는 현상을 막아준다. 세 번째로, 학습률(learning rate)는 gradient decent를 계산할 때, 한 단계씩 손실함수를 변경하는 비율로써, 적정한 학습률을 설정하는 것이 중요하다. 너무 클 경우에는 overshooting하여 발산하는 결과를 가져오고 너무 작으면 local minimum 문제에 있어서 최소값을 찾지 못하는 문제가 발생한다. 네 번째는 각 층의 노드의 수는 한 층에 할당된 노드의 개수를 의미하고 각 층의 개수는 은닉층의 층 개수를 의미한다. 최적화 알고리즘은 파라미터를 수정할 때, 역전파(back-propagation)를 통해서 하게 되는데 이때 알고리즘을 확률적 경사하강법(Stochastic Gradient Descent)[34]이나 Adam 알고리즘[35] 등을 사용한다.

5. 실험환경 및 실험결과

딥뉴럴네트워크를 이용한 침입탐지시스템의 성능을 분석하기 위하여 실험환경으로 텐서플로우(Tensorflow) 머신러닝 라이브러리[36]를 사용하였으며, 서버는 Intel(R) Core(TM) i3-7100 CPU @ 3.90GHz와 GPU는 GeForce GTX 1050을 사용하였다. 실험데이터셋으로는 DARPA에서 제공한 41가지 침입유형이 있는 KDD CUP 99 데이터셋을 활용하여 실험을 하였다. 이러한 실험을 통해서 딥뉴럴네트워크에서 각 파라미터에 의한 성능이 어떻게 되는 지 분석하였다.

5.1 데이터셋

데이터셋은 DARPA 프로그램에 의해서 ID 평가

를 위해 MIT에 있는 Lincon Labs에 의해서 만들어졌다. 이 데이터셋의 주요 목적은 ID의 연구를 분석하고 연구하는 데 사용되었다. 대표적인 데이터셋으로써, 다양한 침입 유형에 포함하여 군사적인 상황과 공공 상황에서도 적용할 수 있는 데이터셋이다. 데이터셋[14]에서 제공되는 10% 증가된 데이터셋을 추가 적용하여 활용하였다. 1999년도에 제안된 KDD CUP 99 [8] 데이터셋은 침입탐지내용이 잘 정의된 버전으로 사용되고 있다. DARPA의 ID 평가는 UNIX 노드 1000개 이상에서 9주 동안 연속적으로 LAN 기반에 공군에서 축척된 네트워크 정보이며, TCP 데이터를 추출하여 training data는 7주, test data는 2주간 나눠서 100명의 사용자로부터 받아서 KDD CUP 99 데이터셋을 생성하였다. MIT 연구실에서 DARPA와 AFLRL에 지원을 받아서 데이터셋을 생성하였으며, 이 데이터셋의 목적은 7가지의 시나리오와 32개의 공격을 조합하여 총 300개의 공격 시뮬레이션을 가능하게 하는 것이다.

침입탐지시스템을 평가하기 위하여 KDD CUP 99는 널리 사용되어 왔으며 41개의 feature로 구성되어 있다. 공격으로 시뮬레이터된 유형들은 DoS(Denial-of-service attack) 공격, U2R 공격(User-to-Root attack), R2L 공격 (Remote-to-local attack), Probing 공격으로 크게 구분된다. DoS은 호스트에 오버헤드가 걸리도록 많은 양의 데이터를 제공하여 정상적인 서비스 제공을 마비시키는 공격 방법이다. 두 번째로 U2R 공격은 사용자의 기존 접근을 통해서 root 권한까지 확장하는 공격 방법이다. 세 번째로 R2L 공격은 권한 없는 사용자가 외부에서 접근 권한을 얻으려고 패킷을 보내는 공격 방법을 의미한다. 네 번째로, Probing 공격은 실제 공격을 하기 전에 시스템의 사전 포트 정보 등을 수집하는 패킷 공격방법을 의미한다.

KDD CUP 99 데이터셋은 3가지 그룹으로 분류된다. 먼저 basic feature는 모든 TCP/IP 네트워크에서 추출할 수 있는 속성들을 의미한다. 두 번째로 traffic feature는 2가지로 구성되어 있는 데, 같은 호스트 기능은 현재 접속과 동일한 목적지를 갖는 호스트여부와 지난 2초 동안 접속만을 조사한 결과, 행동, 서비스 통계결과 등을 보여준다. 반면에 동일한 서비스 기능은 현재 연결과 동일한 서비스를 지닌 2초 동안의 연결만 검사한다. 세 번째로 content features는 로그인 실패 회수, 루트의 접속한 횟수 등에 대한 시스템적인 요소들을 보여준다.

5.2 딥뉴럴네트워크

각 파라미터에 의하여 딥뉴럴네트워크에 대한 성능을 측정하기 위하여 침입을 탐지하는 딥뉴럴네트워크의 구성은 <표 1>과 <표 2>와 같이 각 파라미터와 구조에 따라서 각 탐지 성능을 측정하였다. 딥뉴럴네트워크의 구조에서 각 은닉층의 개수를 DNN 1개 layer에서부터 7개 layer까지 각각의 성능을 테스트한다. 또한, 활성화 함수를 유형 2가지를 각각 적용한 성능 값, 드랍아웃, Epoch을 다르게 하여 침입탐지시스템의 성능을 비교하였다.

〈표 1〉 딥뉴럴네트워크의 구조

각 층의 제 원	수치값
Fully connected layer	1024
Fully connected layer	768
Fully connected layer	512
Fully connected layer	256
Fully connected layer	128
Fully connected layer	64
Fully connected layer	32
Fully connected layer	1
Fully connected layer	ReLU / Sigmoid function

〈표 2〉 딥뉴럴네트워크 파라미터

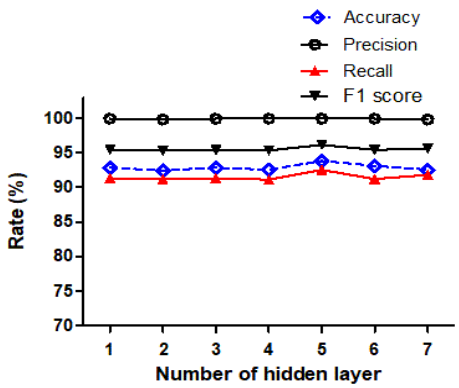
제 원	종류 및 수치값
Optimizer	Adam
Learning rate	0.01
Dropout	0.01 / 0.05 / 0.1
Batch size	64
Epochs	10~100

5.3 성능 측정

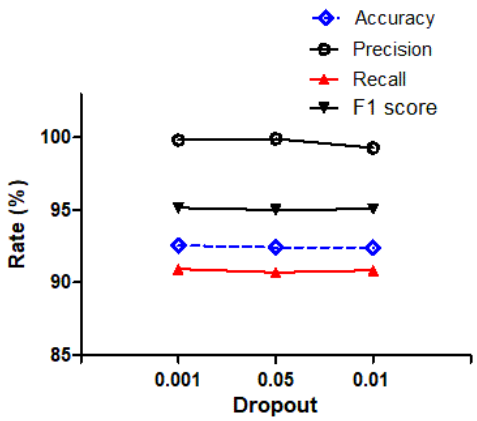
침입탐지시스템에 대한 성능측정은 다음과 같은 요소를 통해서 측정하였다. 정확도(Accuracy)는 전체 레코드 중에서 정확히 분류한 레코드로 확률값을 의미한다. 정밀도(Precision, P)는 TP(True positive)와 FP(False positive)합의 TP의 백분율로 정의가 된다. 재현율(Recall, R)은 TP와 FN(False negative)의 합의 TP의 백분율로 정의된다. F1-score는 정밀도(P)와 재현율의 조화평균(R)을 의미한다.

5.4 실험결과

(그림 2)는 은닉층의 층개수에 따른 딥뉴럴네트워크에 대한 성능을 분석한 결과이다. 결과를 보면 은닉층의 층개수가 많아지더라도 성능의 개선이 이뤄지지 않은 것을 볼 수 있다. 이 실험에서 은닉층의 개수가 5개(5-DNN)일 때 93.1% 정확도의 성능으로 가장 좋은 것을 볼 수 있다. (그림 3)은 드랍아웃(Dropout)에 따른 딥뉴럴네트워크의 성능을 보여준다. 이 때 딥뉴럴네트워크는 은닉층 1개부터 7개(1-DNN~7-DNN)까지의 전체 평균한 값을 보여준다. 이 그림에서 보면 드랍아웃의 확률값이 증가할수록 값 차이는 크지 않지만 약간씩 딥뉴럴네트워크의 성능이 오히려 떨어지는 것을 볼 수 있었다.

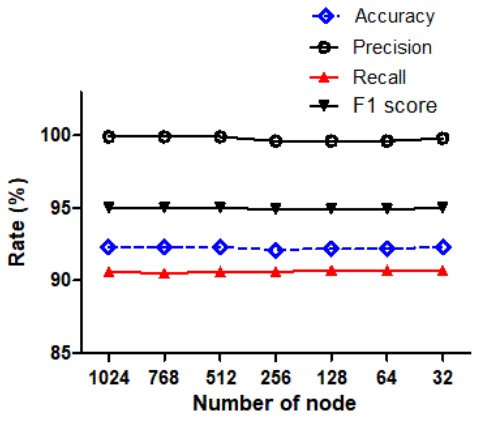


(그림 2) 은닉층 개수에 따른 딥뉴럴네트워크에 대한 성능분석

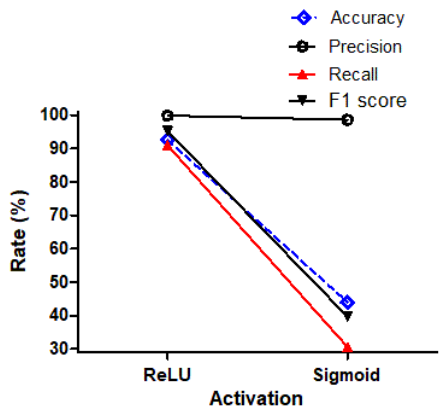


(그림 3) Dropout에 따른 딥뉴럴네트워크에 대한 평균 성능분석

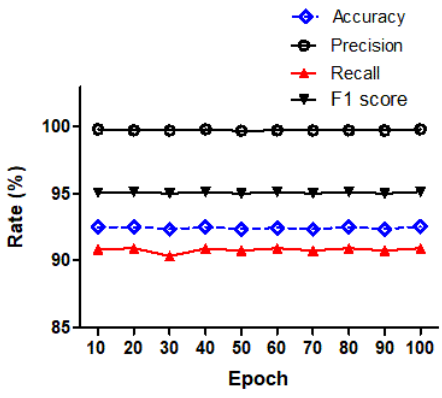
(그림 4)는 은닉층 개수가 1개인 딥뉴럴네트워크(1-DNN)에 노드수에 따른 성능 분석한 결과이다. 결과를 보면 노드수가 1024, 768, 512 일 때, 다른 노드보다 약 1~2% 정확도가 높은 것을 볼 수 있었지만 노드수의 상관없이 거의 성능이 비슷한 것을 볼 수 있었다.



(그림 4) DNN1에서 첫 번째 층에 노드수에 따른 침입탐지시스템의 성능결과



(그림 5) 활성화함수에 따른 침입탐지시스템의 성능결과



(그림 6) Epoch에 따른 침입탐지시스템의 성능결과

(그림 5)는 활성화 함수에 따른 딥뉴럴네트워크의 성능을 분석한 결과이다. 이 때, 딥뉴럴네트워크는 은닉층 1개부터 7개(1-DNN~7-DNN)까지 전체 평균값을 보여준다. 결과를 보면 ReLU 함수를 사용할 때 성능이 더 좋은 것을 볼 수 있었다. 이는 gradient vanish 현상으로 인해서 sigmoid 함수가 취약점이 있지만 ReLU는 특정 임계치 이상부터는 1차원 함수로 증가하기 때문에 gradient vanish 현상을 막아 노드가 잘 학습하는 것을 볼 수 있었다.

(그림 6)은 Epoch 수에 따른 딥뉴럴네트워크의 성능을 분석한 결과이다. 이 때, 딥뉴럴네트워크는 은닉층 1개부터 7개(1-DNN~7-DNN)까지 전체 평균값을 보여준다. Epoch 수가 10 이상일 때, 거의 성능이 비슷하게 유지되는 것을 볼 수 있었다.

<표 3>은 다른 머신들과 비교하여 딥뉴럴네트워크에 대한 성능 분석한 표를 보여준다. 다른 머신러닝기법도 약 92.5~93.8%사이의 성능을 보여주는 것을 볼 수 있었다. 딥뉴럴네트워크에서는 5-DNN일 때 정확도 93.8%와 F1 score가 96.1%으로 가장 성능이 좋은 것을 볼 수 있었다.

〈표 3〉 다른 머신러닝방법과의 성능 비교 - LR은 Lear Regression을 의미하고, NB는 Naive baysian을 의미, KNN는 K-Nearest Neighbors를 의미, DT는 Decision Tree를 의미, Adaboost는 Adaptive Boosting을 의미, RF는 Random Forest를 의미함.

제 원	Accuracy	Precision	Recall	F1 score
1-DNN	92.80%	99.90%	91.30%	95.40%
2-DNN	92.40%	99.80%	91.20%	95.30%
3-DNN	92.80%	99.90%	91.30%	95.40%
4-DNN	92.50%	99.90%	91.10%	95.30%
5-DNN	93.80%	99.90%	92.50%	96.10%
6-DNN	93.10%	99.90%	91.20%	95.40%
7-DNN	92.50%	99.80%	91.80%	95.60%
LR	84.90%	98.90%	82.10%	89.70%
NB	92.90%	98.90%	92.40%	95.50%
KNN	92.90%	99.80%	91.50%	95.50%
DT	92.90%	99.90%	91.30%	95.40%
Adaboost	92.50%	99.60%	91.40%	95.30%
RF	92.60%	99.90%	91.10%	95.30%

6. 토론

침입탐지를 위한 딥뉴럴네트워크를 구성할 때, 딥뉴럴네트워크에 대한 파라미터에 따른 성능 변화에 대해서 분석할 필요성이 있다. 왜냐하면 잘못된 파라미터 설정으로 시스템의 성능에 영향을 미칠 수 있기 때문이다. 따라서 이 장에서는 각 파리미터에 대한 성능 분석을 정리하였다.

딥뉴럴네트워크에 있는 파라미터를 조정하여 성능을 분석한 것을 보면, 상대적으로 은닉층의 개수, 각 층의 노드의 수, Dropout에 따른 성능이 거의 유사한 것을 볼 수 있었다. 또한 Epoch 수도 10이상일 경우에는 비슷한 성능으로 10 epoch이면 거의 최적화된 파라미터로 설정된 것을 볼 수 있었다.

반면에 활성화함수 선정은 중요한 것을 볼 수 있었다. LeRU를 사용하는 대신 Sigmoid function을 사용하게 되면 딥뉴럴네트워크의 성능이 떨어지는 것을 볼 수 있었다. gradient vanish 현상을 줄이는 활성화 함수의 선정이 중요한 것을 볼 수 있었다.

이 본문에서 은닉층의 개수가 5개인 딥뉴럴네트워크(5-DNN)일 때 성능이 좋은 것을 볼 수 있었다. 딥뉴럴네트워크의 구조 선정에 있어서 약간씩 성능을 달라지는 것을 볼 수 있었지만 최적의 성능을 찾기 위해서 여러 가지 파라미터를 조정하고 테스트하면서 찾는 과정이 요구되는 것을 볼 수 있었다.

7. 결론

이 논문에서는 딥뉴럴네트워크를 이용하여 침입탐지시스템에 대한 성능 분석을 각 파라미터를 변경하여 분석을 하였다. 분석 결과를 살펴보면 은닉층의 개수, Epoch나 각 노드의 수가 증가할수록 더 이상의 성능이 개선되는 것은 없는 것을 볼 수 있었다. 반면에 활성화 함수가 Sigmoid function 대신에 LeRU 일 때 성능이 좋은 것을 볼 수 있었다. 다른 머신러닝 방법과 비교하였을 때, 은닉층의 개수가 5개 일 때, 93.8% 정확도와 96.1% F1 score를 갖는 것을 볼 수 있었다.

향후 연구로는 네트워크에 활용되는 데이터셋인 NSL-KDD 데이터 등으로 확장하여 실험할 수 있다. 침입탐지 분야에 있어서 잘못탐지(false detection)을 줄이는 체계적인 분석도 흥미로운 향후 주제가 될 것이다. 또한 Generative adversarial network (적대적 생성 네트워크)를 이용하여 악의적인 코드를 생성하거나 침입탐지하는 방법에 대하여 분석할 예정이다.

Acknowledgement

본 연구는 과학기술정보통신부의 재원으로 정보통신기획평가원(IITP)의 지원(No.2018-0-00420, No.2019-0-00273)을 받아 수행된 연구임.

참고문헌

〔1〕 Tidjon, Lionel N., Marc Frappier, and Amel Mammar. "Intrusion Detection Systems: A Cross-Domain Overview." IEEE Communications Surveys & Tutorials (2019).

〔2〕 Mawla, Tanjila, Sharmishtha Dutta, and Md Forhad Rabbi. "Temporal Signature Mining for Network Intrusion Detection Using TEMR." Emerging Technologies in Data Mining and Information Security. Springer, Singapore, 2019. 645-655.

〔3〕 Alhakami, Wajdi, et al. "Network Anomaly Intrusion Detection Using a Nonparametric Bayesian Approach and Feature Selection." IEEE Access 7 (2019): 52181-52190.

〔4〕 Schmidhuber, Jürgen. "Deep learning in neural networks: An overview." Neural networks 61 (2015): 85-117.

〔5〕 Gao, Shangqi, and Xiahai Zhuang. "Multi-scale deep neural networks for real image super-resolution." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2019.

〔6〕 Huang, Kun-Yi, et al. "Speech Emotion Recognition Using Deep Neural Network Considering Verbal and Nonverbal Speech Sounds." ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019.

〔7〕 Akhundov, Riad, et al. "Development of a deep neural network for automated electromyographic pattern classification." Journal of Experimental Biology 222.5 (2019): jeb198101.

〔8〕 Tavallaee, Mahbod, et al. "A detailed analysis of the KDD CUP 99 data set." 2009 IEEE Symposium on Computational Intelligence for

Security and Defense Applications. IEEE, 2009.

〔9〕 Wang, HaiYing, Min Yang, and John Stufken. "Information-based optimal subdata selection for big data linear regression." Journal of the American Statistical Association 114.525 (2019): 393-405.

〔10〕 Chen, Jiangning, et al. "Naive Bayes with Correlation Factor for Text Classification Problem." arXiv preprint arXiv:1905.06115 (2019).

〔11〕 Zheng, Wei, et al. "An Improved k-Nearest Neighbor Classification Algorithm Using Shared Nearest Neighbor Similarity." Metallurgical & Mining Industry 10 (2015).

〔12〕 Song, Yan-Yan, and L. U. Ying. "Decision tree methods: applications for classification and prediction." Shanghai archives of psychiatry 27.2 (2015): 130.

〔13〕 Farnaaz, Nabila, and M. A. Jabbar. "Random forest modeling for network intrusion detection system." Procedia Computer Science 89 (2016): 213-217.

〔14〕 http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

〔15〕 Kwon, Hyun, et al. "Optimal Cluster Expansion-Based Intrusion Tolerant System to Prevent Denial of Service Attacks." Applied Sciences 7.11 (2017): 1186.2

〔16〕 KBennett, Kristin P., and Ayhan Demiriz. "Semi-supervised support vector machines." Advances in Neural Information processing systems. 1999.

〔17〕 D. Shin, K. Choi, S. Chune and H. Choi, Malicious Traffic Detection Using K-means, The Journal of Korean Institute of Communications and Information Sciences, 41(2), pp. 277-284. 2016.

〔18〕 M. Tahir, W. Hassan, A. Md Said, N. Zakaria, N. Katuk, N. Kabir, M. Omar, O. Ghazali and N. Yahya, Hybrid machine learning technique for intrusion detection system, 5th International Conference on Computing and Informatics (ICOCI), 2015.

〔19〕 N. Gao, L. Gao, Q. Gao, and H. Wang An Intrusion Detection Model Based on Deep Belief Networks, Advanced Cloud and Big Data (CBD), 2014 Second International Conference on, pp. 247-252, 2014.

〔20〕 S. Jo, H. Sung, and B. Ahn A Comparative Study on the Performance of SVM and an Artificial Neural Network in Intrusion Detection, Journal of the Korea Academia-Industrial cooperation Society, 17(2), pp. 703- 711, 2016.

〔21〕 Kim, Kwangjo, Muhamad Erza Aminanto, and Harry Chandra Tanuwidjaja. Network Intrusion Detection Using Deep Learning: A Feature Learning Approach. Springer, 2018.

〔22〕 Mathai, K. James. "Performance Comparison of Intrusion Detection System Between Deep Belief Network (DBN) Algorithm and State Preserving Extreme Learning Machine (SPELM) Algorithm." 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT). IEEE, 2019.

〔23〕 Aggarwal, Preeti, and Deepak Dahiya. "Contribution of four class labeled attributes of kdd dataset on detection and false alarm rate for intrusion detection system." Indian Journal of Science and Technology 9.5 (2016): 1-8.

〔24〕 Gurung, Sandeep, Mirnal Kanti Ghose, and Aroj Subedi. "Deep learning approach on network intrusion detection system using

NSL-KDD dataset." International Journal of Computer Network and Information Security (IJCNIS) 11.3 (2019): 8-14.

〔25〕 Özgür, Atilla, and Hamit Erdem. "A review of KDD99 dataset usage in intrusion detection and machine learning between 2010 and 2015." PeerJ Preprints 4 (2016): e1954v1.

〔26〕 Meena, Gaurav, and Ravi Raj Choudhary. "A review paper on IDS classification using KDD 99 and NSL KDD dataset in WEKA." 2017 International Conference on Computer, Communications and Electronics (Comptelix). IEEE, 2017.

〔27〕 O. AI-Jarrah and A. Arafat Network intrusion detection system using neural network classification of attack behavior, Journal of Advances in Information Technology Vol, 6(1), 2015.

〔28〕 Li, Yuanzhi, and Yang Yuan. "Convergence analysis of two-layer neural networks with relu activation." Advances in Neural Information Processing Systems. 2017.

〔29〕 와셈, et al. "컨볼루션 뉴럴 네트워크를 이용한 군중 행동 감지." 한국차세대컴퓨팅학회 논문지 15.6: 7-14.

〔30〕 정혜진, and 나연묵. "데이터 센터의 리소스를 효율적으로 관리하기 위한 인공지능 기반의 프레임워크." 한국차세대컴퓨팅학회 논문지 15.2: 79-88.

〔31〕 김준태, and 배창석. "드라이빙 게임 환경에서의 효과적인 학습 데이터 수집 및 컨볼루셔널 뉴럴 네트워크 기반의 자율주행." 한국차세대컴퓨팅학회 논문지 15.1 (2019): 28-37.

〔32〕 Vigneswaran, K. Rahul, et al. "Evaluating shallow and deep neural networks for network intrusion detection systems in cyber security." 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, 2018.

〔33〕 Yin, Xinyou, et al. "A flexible sigmoid function of determinate growth." Annals of botany 91.3 (2003): 361-371.

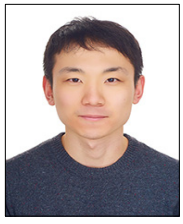
〔34〕 Bottou, Léon. "Large-scale machine learning with stochastic gradient descent." Proceedings of COMPSTAT'2010. Physica-Verlag HD, 2010. 177-186.

〔35〕 Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014).

〔36〕 Abadi, Martín, et al. "Tensorflow: A system for large-scale machine learning." 12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16). 2016.

저자소개

◆ 권현



- 2010년 육군사관학교 수학과 졸업(학사)
- 2015년 KAIST 전산학부 졸업(석사)
- 2015년~2016년 육군사관학교 전자공학과
- 2020년 KAIST 전산학부 졸업(박사)
- 2020년~현재 육군사관학교 전자공학과 조교수
- 관심분야: 머신러닝, 인공지능 보안, 시스템 보안, 침입감내시스템

◆ 방승호

- 2006년 육군사관학교 군사학 졸업(학사)
- 2015년 KAIST 산업공학과 졸업(석사)
- 2019년~현재 합동참모본부 전력기획부
- 관심분야: M&S, 인공지능

◆ 박기웅

- 2005년 연세대학교 컴퓨터공학과 졸업 (학사)
- 2007년 KAIST 전자전산학부 졸업(석사)
- 2012년 KAIST 전기및전자 졸업(박사)
- 2008년 Microsoft Research Asia, Research Intern
- 2009년 Microsoft Research, Graduate Research Fellow
- 2012년 국가보안기술연구소 연구원
- 2012년~2016년 대전대학교 정보보안학과 교수
- 2016년~현재 세종대학교 정보보호학화 교수
- 관심분야: 시스템보안, 모바일 및 클라우드 컴퓨팅, 보안프로토콜 등