Journal of KIIT. Vol. 22, No. 7, pp. 45-54, Jul. 31, 2024. pISSN 1598-8619, eISSN 2093-7571
http://dx.doi.org/10.14801/jkiit.2024.22.7.45

네트워크 침입 탐지 분야에서 공격 유형 분류를 위한 합성곱 신경망 모델에 대한 연구

주승세*[1], 이성혁*[2], 유재학**[1], 문대성**[2], 배지훈***

Study on a Convolutional Neural Network Model for Attack Type Classification in the Field of Network Intrusion Detection

SeungSae Joo*[1], SungHyuk Lee*[2], Jaehak Yu**[1], Daesung Moon**[2], and Ji-Hoon Bae***

이 논문은 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임
(00235509, ICT융합 공공 서비스・인프라의 암호화 사이버위협에 대한 네트워크 행위기반 보안관제 기술 개발)

요 약

네트워크 상에서 발생하는 사이버 공격으로 인한 피해가 매년 증가하고 있다. 기존에는 알려진 공격 유형을 탐지하기 위한 지도 학습 기반의 딥러닝 모델과, 비정상 상황을 탐지하기 위한 비지도 학습 기반의 딥러닝 모델에 대한 연구가 주로 이루어졌다. 하지만 이러한 접근 방식은 새로운 유형의 공격에 대한 정확한 탐지와, 적절한 대응이 어렵다는 한계가 존재한다. 본 논문에서는 새로운 유형의 공격을 탐지하기 위한 전이 학습 기반 딥러닝 모델을 제안한다. 이는 두 단계의 학습 과정으로 구성되어 있다. 단계 1에서는 Bot-IoT 소스 도메인 데이터 셋으로부터 다양한 특징을 추출하고, 이를 기반으로 단계 2에서는 UNSW-NB15 타겟 도메인 데이터 셋으로 재학습을 수행한다. 다양한 공격 유형에 대하여 제안 모델은 약 94.21%의 분류 정확도를 보여주었으며, 이는 기존 합성곱 신경망 기반 모델 보다 약 1.84% 향상된 결과를 제공하는 것을 실험적으로 관찰할 수 있었다.

Abstract

Damage from cyber attacks on networks is increasing every year. Previously, research has mainly focused on supervised learning-based deep learning models for detecting known attack types and unsupervised learning-based models for detecting anomalies. However, these approaches have limitations in accurately detecting new types of attacks and responding appropriately. In this paper, we propose a transfer learning-based deep learning model for detecting new types of attacks. It consists of two stages of training. In Stage 1, various features are extracted from the Bot-IoT source domain dataset, and based on this, Stage 2 performs retraining on the UNSW-NB15 target domain dataset. For various attack types, the proposed model showed approximately 94.21% classification accuracy, which experimentally observed to be about 1.84% higher than that of existing convolutional neural network-based models.

Keywords

network intrusion detection, transfer learning, attack type classification, UNSW-NB15

* 대구가톨릭대학교 AI빅데이터공학과
- ORCID[1]: https://orcid.org/0009-0001-1623-5424
- ORCID[2]: https://orcid.org/0009-0009-5295-0204
** 한국전자통신연구원 선임연구원
- ORCID[1]: https://orcid.org/0000-0003-0281-7666
- ORCID[2]: https://orcid.org/0000-0002-9961-0401
*** 대구가톨릭대학교 AI빅데이터공학과 조교수(교신저자)
- ORCID: http://orcid.org/0000-0002-0035-5261

・Received: Apr. 29, 2024, Revised: Jun. 20, 2024, Accepted: Jun. 23, 2024
・Corresponding Author: Ji-Hoon Bae
Dept. of AI and Big Data Engineering, Hayang-ro 13-13, Hayang-eup, Gyeongsan-si, Gyeongsangbuk-do, Korea
Tel.: +82-53-850-2883, Email: jihbae@cu.ac.kr

Ⅰ. 서 론

현재, 사물 인터넷(Internet of things)을 통한 다양한 기기의 등장으로 네트워크에서 송수신되는 트래픽의 양과 종류가 다양해지고 있다[1]. 또한, 인터넷 기술의 발전을 통해 용이한 정보 획득과 다양한 산업 분야에서의 업무 효율성 향상, 비용 절감 등의 성과를 이루어내고 있다. 하지만 이는 개인정보 유출, 사이버 테러, 정보 도난, 물리적 시스템 작동 불능 등을 목적으로 하는 악의적인 사이버 공격에 취약하다는 단점이 있다. 또한, 사이버 공격 기법의 다양화와 지능화로 인해 사이버 범죄에 대한 피해가 증가하고 있으며, 이로 인해 공격 유형에 적합한 대응이 요구되고 있다[2].

기존에는 네트워크 침입 탐지를 위해 지도 학습 기반의 연구가 이루어졌다[3]. 하지만, 지도 학습 기반 모델은 공격 탐지에는 효과적이나 새로운 유형의 공격에 대응하기 어렵다는 문제점이 있다[4]. 또한, 모델을 최신 상태로 유지하기 위해서는 지속적인 학습 데이터가 필요하며, 이는 비용과 시간이 많이 소요되는 제약이 있다[5]. 이러한 문제를 해결하기 위해 정상의 특징 학습을 통해 네트워크 공격을 탐지하는 비지도 학습 기반 모델 연구가 활발히 진행되고 있지만, 이러한 방법은 공격 유형에 대한 예측이 어려워 유형에 대한 적절한 대응을 제공하기 어렵다는 한계가 있다[6]-[8].

따라서, 본 논문에서는 기존 네트워크 침입 탐지 연구의 문제점을 해결하기 위해 소량의 데이터만으로도 높은 성능을 달성할 수 있고, 다양한 환경에 적용 가능한 전이 학습 기반 네트워크 침입 탐지 모델을 제안한다.

본 논문의 2장에서는 모델 학습에 사용된 데이터와 수행된 전처리 기법에 대하여 설명하고, 3장에서는 새로운 사이버 공격 유형을 탐지하기 위한 전이 학습 기반 네트워크 침입 탐지 모델을 소개한다. 4장에서는 실험 결과를 분석하며, 마지막으로 5장에서는 본 논문의 결론과 향후 연구 계획을 제시하여 마무리한다.

Ⅱ. 관련 연구

최근, 머신러닝과 딥러닝 같은 인공지능 기술이 다양한 산업 분야에 적용됨에 따라, 인공지능 기술 기반의 네트워크 침입 탐지 시스템에 대한 요구가 증가하고 있다.

기존에는 공격 유형을 정확하게 분류하기 위한 지도 학습 기반 딥러닝 모델에 대한 연구가 주로 수행되었다. S. W. Lim et al.[9]은 의사 결정 트리, Random Forest, Extra Tree, XGBoost, LightGBM과 같은 대표적인 기계학습 모델을 적용한 네트워크 침입 탐지 기법을 연구하였으며, 실험을 통해 XGBoost 기법이 가장 정확하게 공격 유형을 분류함을 보여주었다. L. Ashiku et al.[10]은 합성곱 신경망 기반의 네트워크 침입 탐지 시스템을 제안하였으며, 약 95.6%의 정확도를 보였다. 하지만, 이러한 지도 학습 기반의 네트워크 침입 탐지 모델은 새로운 유형의 공격이 발생할 경우 정확한 탐지가 어려운 한계가 있다.

이러한 문제를 해결하기 위해 이상 탐지 기법을 적용한 네트워크 침입 탐지 시스템에 대한 연구가 활발히 진행되고 있다. K. H. Kang[11]은 네트워크 침입 탐지를 위한 비지도 학습 기반 오토인코더(Auto-encoder) 모델을 제안하였으며, 약 88.9%의 정확도를 보였다. 여기서, 오토인코더 모델은 정상 데이터로 구성된 학습 데이터를 압축 및 복원하며, 입력과 출력의 차이를 최소화하기 위하여 학습한다. 복원 성능을 기준으로 정상 범위를 설정하고, 이 범위 내에 분포하는 데이터는 정상으로, 범위를 벗어나는 데이터는 비정상으로 분류한다. 이러한 비지도 학습 기반 모델은 비정상적인 상황을 탐지할 수 있지만, 공격 유형을 파악할 수 없다. 사이버 공격은 공격 유형에 따라 대응 방법이 다르기 때문에, 비지도 학습 기반 모델은 공격 유형에 따른 적절한 대응 방법을 적용하기 어렵다는 한계가 존재한다.

따라서, 본 논문에서는 소량의 데이터로도 기존에 알려지지 않은 공격 유형을 탐지할 수 있는 전이 학습 기법 기반의 딥러닝 모델을 제안한다.

Ⅲ. 데이터 수집 및 전처리

본 논문에서는 소스 도메인(Source domain)과 타

겟 도메인(Target domain) 간의 전이 학습 기법 적용을 위해 Bot-IoT와 UNSW-NB15 두 가지 유형의 데이터 셋을 고려한다. 먼저, 소스 도메인 데이터로 설정한 Bot-IoT는 UNSW Canberra에서 기존 데이터의 부족한 사이버 공격 정보를 보완하기 위해 현실적인 테스트 베드 환경에서 생성된 데이터 셋으로, 봇넷(Botnet) 공격에 대한 데이터 셋이다[12].

해당 데이터 셋은 46개의 특성과 7,200만 개 이상의 레코드로 구성되어 있으며, 99% 이상의 봇넷 트래픽과 1% 미만의 정상 트래픽으로 구성되어 있다. 여기서 봇넷 트래픽은 4가지 공격 유형(DoS, DDoS, Theft, Reconnaissance)에 대한 트래픽 데이터로 구성된다.

다음으로, 타겟 도메인 데이터로 설정된 UNSW-NB15은 UNSW Canberra에서 'IXIA PerfectStorm tool'을 사용하여 네트워크 패킷을 수집하고, 'Argus'와 'Bro-IDS Tool'을 통해 정상과 공격 트래픽을 분류하여 생성된 데이터 셋이다. 해당 데이터 셋은 49개의 특성과 200만 개 이상의 레코드로 구성되어 있으며, 실제 네트워크상에서 발생하는 9가지 공격 유형(DoS, Exploits, Generic, Reconnaissance, Analysis, Backdoor, Fuzzers, Shellcode, Worms)에 대한 트래픽과 정상 트래픽으로 구성되어 있다[13].

상기 기술된 소스 및 타겟의 데이터 셋은 심각한 클래스 불균형을 가지고 있다. 여기서, 클래스 불균형은 소수 클래스에 속하는 데이터에 비해 다수 클래스에 속하는 데이터가 과도하게 분포하는 것으로, 다수 클래스가 소수 클래스의 영역을 침범하여 모델의 성능을 저하시키는 문제를 의미한다[14]. 이러한 문제를 해결하기 위해 주로 오버 샘플링(Oversampling) 기법과 언더 샘플링(Undersampling) 기법이 활용되며, 오버 샘플링 기법은 소수 클래스 데이터 수를 증가시키는 기법을, 언더 샘플링 기법은 다수 클래스 데이터의 일부만 사용하는 기법을 각각 의미한다. 해당 기법은 다수 클래스와 소수 클래스의 데이터 비율을 축소하기 위해 수행되며, 일반적으로 클래스 균형화는 각 클래스의 데이터 비율이 1:3을 넘지 않도록 설정한다[15].

본 연구에서는 전이 학습에 활용되는 소스 도메인 모델 학습을 위해, BoT-IoT 데이터 셋에서 클래스 불균형을 최소화하고자 데이터 개수가 10,000개 미만인 'Normal'과 'Theft' 클래스를 제외하여 그림 1과 같이 데이터 셋을 구성하였다. 또한, 타겟 도메인의 UNSW-NB15 데이터 셋은 그림 1의 Bot-IoT와 동일한 3개의 공격 유형과 새로운 'fuzzers' 공격 유형을 그림 2와 같이 구성하였다. 이때, 클래스 불균형 문제를 완화하고 전이 학습 효과를 검증하기 위해, 언더 샘플링 기법을 적용하여 각 클래스의 데이터 개수의 비율이 1:2을 넘지 않도록 설정하였다. 이를 통해 타겟 도메인에서의 전이 학습 모델의 효과를 검증하도록 한다.

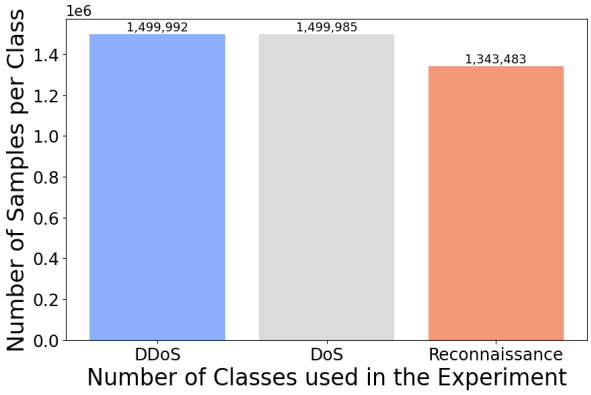

그림 1. 소스 도메인(BoT-IoT) 학습 데이터 구성
Fig. 1. Source domain (BoT-IoT)) training data configuration

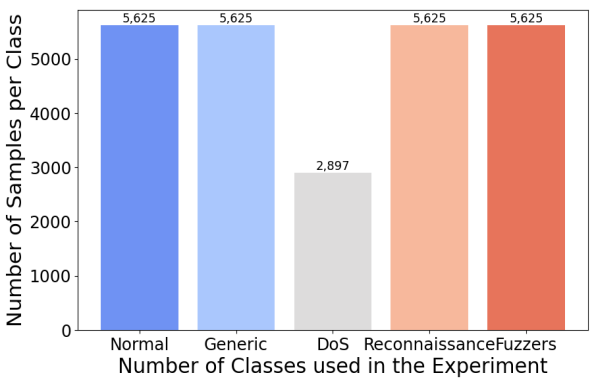

그림 2. 타겟 도메인(UNSW-NB15) 학습 데이터 구성
Fig. 2. Target domain (UNSW-NB15) training data configuration

네트워크 트래픽 데이터는 일반적으로 명목형, 정수, 부동 소수점, 타임스탬프와 같이 이진수로 이루어진 다양한 특성들로 구성되어 있다[16]. 따라서, 본 연구에서는 딥러닝 모델 학습을 위해 명목형 특

성 데이터에 원-핫 인코딩(One-hot encoding)을 적용하여 명목형 데이터를 수치형 데이터로 변환한다[17]. 그 후, 값이 0에 집중된 특성들은 로그 함수를 적용하여 스케일을 조정하며[18], 해당 기법을 적용한 데이터 형태는 그림 3과 같다. 마지막으로, 모든 특성에 대하여 각각 평균이 0, 표준편차가 1이 되도록 식 (1)의 정규화 기법을 적용하여 모델 학습에 활용될 데이터를 구성한다.

$$x_{norm} = \frac{x - \mu}{\sigma} \tag{1}$$

여기서, x는 원본 데이터를, μ는 x의 평균을, σ는 x의 표준편차를 각각 의미한다. x_{norm}은 상기에 기술한 값들을 활용한 정규화된 특성 값을 의미한다.

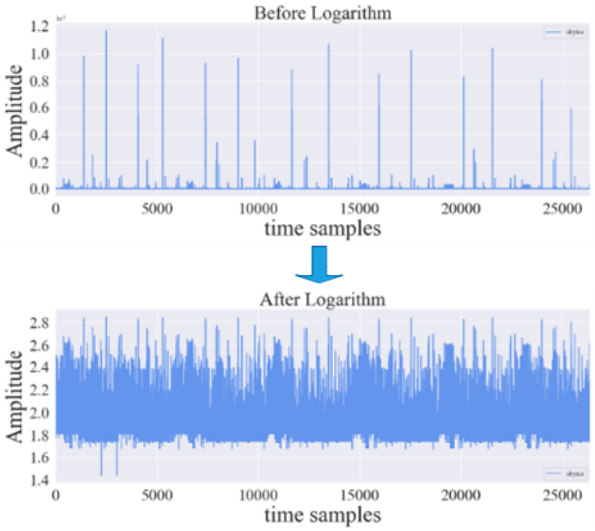

그림 3. 로그 스케일 기법을 적용한 데이터 전처리 결과
Fig. 3. Data preprocessing results applying log scale technique

Ⅳ. 전이 학습 기반 네트워크 침입 탐지 모델

사이버 공격은 다양한 유형으로 네트워크에서 발생하며, 이러한 공격의 특징을 수동으로 감지하기에는 어려움이 있다[19]. 일반적으로 새로운 공격 유형에 대한 데이터는 정상 데이터보다 수집이 매우 제한적으로 이루어진다. 이러한 이유로 데이터 부족 문제가 쉽게 발생하며, 이는 소규모 데이터 셋에서 딥러닝 모델을 학습할 때 과적합(Overfitting)의 원인이 된다[20]. 전이 학습은 사전 학습된 딥러닝 모델이 추출한 특징을 이전받아 새로운 데이터 셋으로 재학습을 수행하는 기법으로, 소량의 데이터만으로도 빠른 학습 속도와 정확한 분류 정확도를 보여주는 것이 입증되었다[21]. 따라서, 본 논문에서는 소량의 데이터로도 높은 분류 정확도를 제공할 수 있는 전이 학습 기반의 네트워크 공격 유형 분류 모델을 제안한다.

본 논문에서 제안하는 네트워크 공격 유형 분류 모델은 먼저 소스 도메인의 Bot-IoT 데이터 셋에 1차원 합성곱 신경망(One-Dimensional Convolutional Neural Network, 1D CNN) 모델을 적용하여 고차원 특징 맵을 추출한다[22]. 이후에는 소스 도메인에는 존재하지 않는 새로운 공격 유형을 포함한 다양한 공격 유형들을 분류하기 위해 밀집 연결 층(Dense layer)으로 구성된 새로운 분류기를 도입하여 전이 학습을 수행한다. 그런 다음, 사전 학습된 모델의 분류기를 제거하고 새로운 분류기를 추가하여 타겟 도메인의 UNSW-NB15 데이터 셋으로 모델을 재학습하며, 새로운 공격에 대한 데이터가 소량이라도 정확하게 분류할 수 있도록 총 두 단계에 걸쳐 학습을 진행한다.

먼저, 그림 4의 단계 1(stage 1)은 데이터로부터 유용한 특징을 효과적으로 추출하기 위해 1차원 합성곱 연산을 활용한다. 이를 통해 Bot-IoT 데이터 셋으로부터 합성곱 출력 특징을 추출하기 위한 모델을 설계하고 학습을 진행하였다. 그림 4에서 첫 번째 1차원 합성곱 층은 10×1 커널 필터를 적용하였으며, 가장자리 값들에 대한 정보 손실을 줄이기 위해 합성곱 층의 입출력 크기를 동일하게 생성하는 zero-padding 기법을 적용하여 32개의 2차원 특징맵을 추출하였다. 다음으로, 두 번째 1차원 합성곱 층 또한 동일한 기법을 적용하여 64개의 2차원 특징맵을 추출하였다. 이때, 모든 1차원 합성곱 층의 활성화 함수는 기울기 소실 문제를 줄이기 위하여 ReLU(Rectified Linear Unit)를 적용하였다[23]. 또한, 합성곱 층 이후에 최댓값 풀링 연산 층을 추가하여 모델 구조의 복잡도를 점진적으로 줄이도록 한다. 이후, 평탄화(Flatten) 층을 추가하여 1차원의 특징 벡터를 생성하고, 이를 밀집 연결 층으로 구성된 분류기에 입력하여 공격 유형에 대한 분류를 수행한다[24].

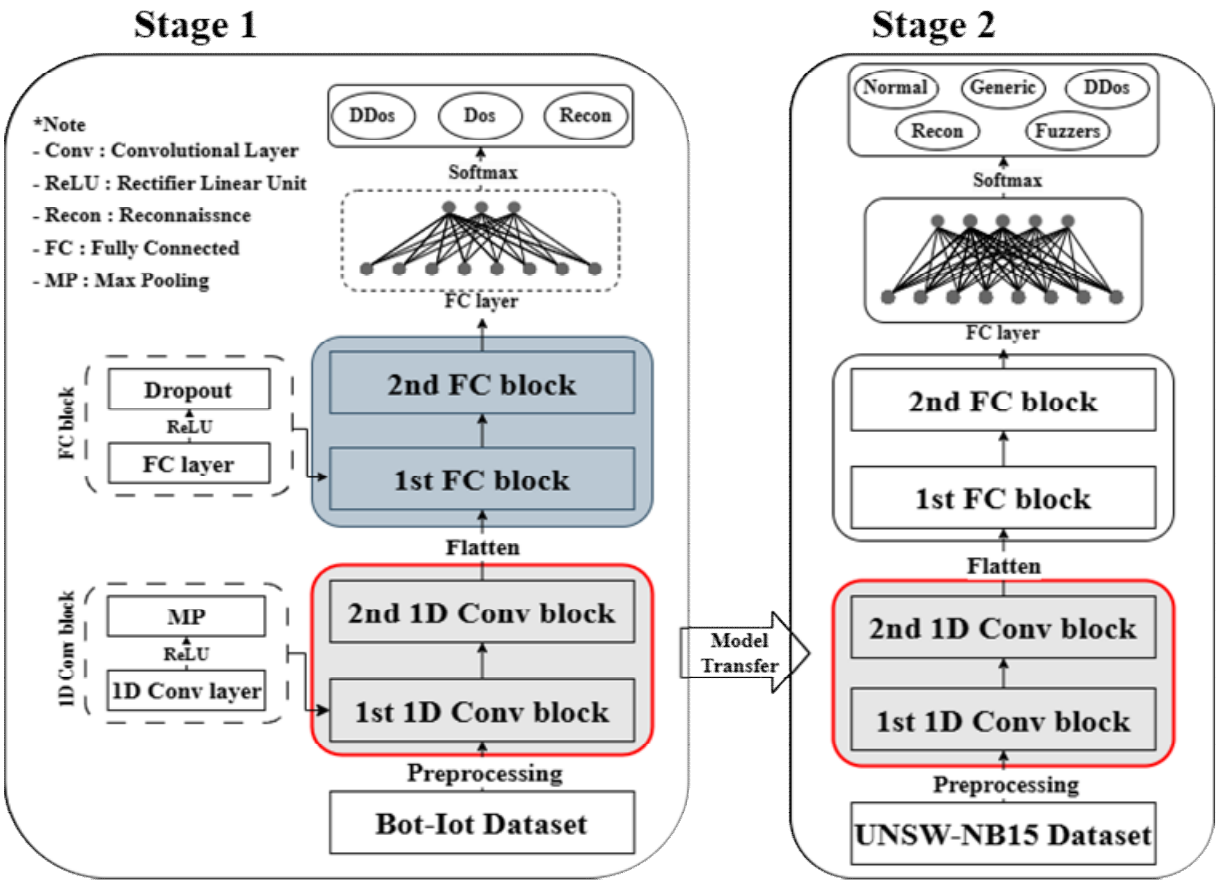

그림 4. 전이 학습 기반의 네트워크 침입 탐지 모델 구조
Fig. 4. Transfer learning-based network intrusion detection model structure

그림 1의 3개의 공격 유형 패턴의 특징을 학습하기 위하여 분류기의 마지막 출력 층의 뉴런은 3개로 구성되며, 소프트맥스(Softmax) 활성화 함수를 사용한다.

그림 4의 단계 1 학습이 완료되면, 사전 학습된 모델의 분류기를 제거한 후, 이를 그림 4의 단계 2(Stage 2)로 모델 이전을 수행한다. 그리고, 이전된 모델에 새로운 분류기를 추가한다. 해당 모델은 사전에 학습한 공격 유형과 학습하지 않은 새로운 공격 유형, 그리고, 정상에 대한 소량의 데이터로 재학습함으로써, 새로운 공격 유형을 소량의 데이터만으로도 정확하게 탐지할 수 있도록 한다.

V. 실험 결과

본 논문에서는 정상 및 공격 데이터를 각각 7,500개, 26,363개를 사용하였다. 여기서, 공격 데이터는 Generic, DoS, Reconnaissance와 Fuzzers 클래스에 대한 데이터를 의미한다. 네트워크 침입 탐지 모델의 성능 평가를 위해 훈련 및 테스트 셋을 7.5:2.5 비율로 분할하여 구성하였으며, 이에 대한 정확한 데이터 개수는 그림 2와 같이 주어진다. 또한, 모델 학습을 위한 파라미터 설정은 배치 사이즈 32, 에포크 100으로 설정하였으며, 학습 알고리즘으로는 Adam(Adaptive moment estimation)을 활용하였다[25]. 모델 학습 시, 학습률 스케줄링(Learning rate scheduling) 기법을 적용하여 초기에 5e-4로 시작하여, 60 에포크 이후로는 5e-5, 80 에포크 이후로는 5e-6으로 점진적으로 학습률을 감소시키며 학습을 진행하였다[26]. 상기 설정한 실험 조건에 따른 전이 학습 기반 모델 학습 결과 그림 5 및 6와 같이 학습 횟수가 진행됨에 따라 안정적으로 손실은 감소하고 정확도 성능은 향상되는 것을 확인할 수 있다.

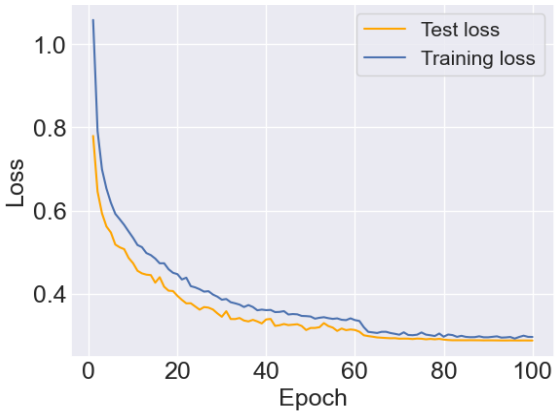

그림 5. 제안된 모델의 학습 및 검증 손실함수 성능 그래프
Fig. 5. Training and validation loss function performance graph of the proposed model

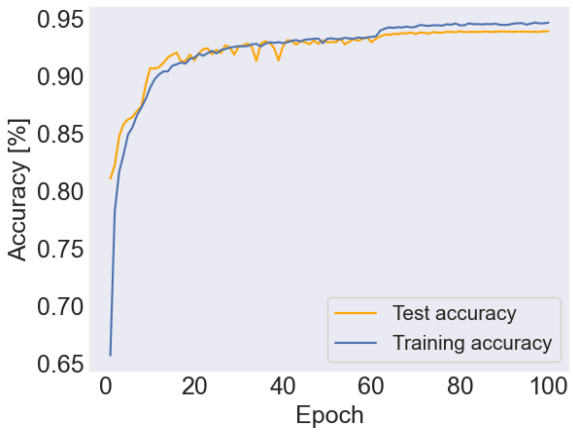

그림 6. 제안된 모델의 학습 및 검증 정확도 성능 그래프
Fig. 6. Training and validation accuracy performance graph of the proposed model

다음으로, 본 논문에서는 전이 학습을 기반으로 한 네트워크 침입 탐지 모델과 단계 2의 타겟 도메인 모델 구조, 동일한 구조의 전이 학습을 적용하지 않은 모델, 로지스틱 회귀와 SVM(Support Vector Machine)의 성능을 상호 비교하고 분석하고자 한다. 모델의 성능을 평가하기 위한 측정 지표는 혼동행렬을 활용하며, 그림 7과 그림 8은 비교 모델 중 가장 우수한 성능을 보인 모델과 제안된 모델의 평가 결과이다.

혼동행렬의 대각 성분은 모델이 정확하게 예측한 데이터의 수를 의미한다. 상기 그림 7, 8을 통하여 알 수 있듯, 전이 학습 기법을 적용한 모델이 전이 학습 기법을 적용하지 않은 모델보다 전반적으로 우수한 성능을 보여주었다. 특히, 새로운 공격 유형에 대한 탐입 탐지 정확도는 다른 공격 유형에 비해 월등히 향상된 성능을 보이는 것을 확인할 수 있다.

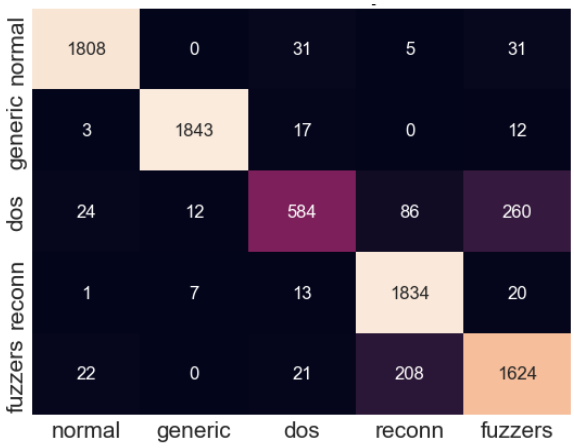

그림 7. 전이 학습 기법을 적용하지 않은 모델의 혼동행렬
Fig. 7. Confusion matrices in the model that do not apply transfer learning techniques

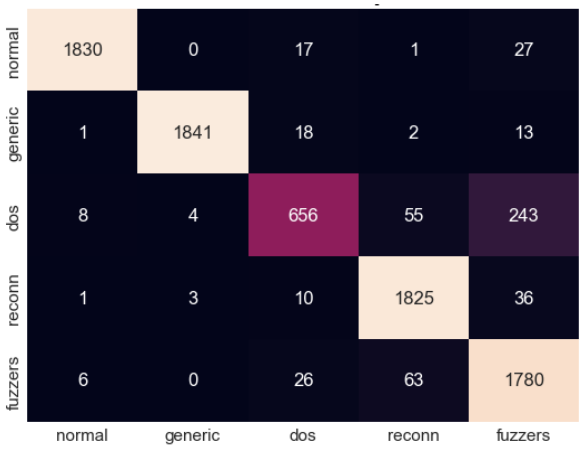

그림 8. 전이 학습 기법을 적용한 모델의 혼동행렬
Fig. 8. Confusion matrix in the model using transfer learning techniques

또한, 모델의 예측 결과와 실제 레이블 간의 관계를 수치화하기 위해 머신러닝 분야에서 모델의 성능 평가 지표로 널리 활용되고 있는 정확도(Accuracy), 정밀도(Precision), 재현율(Recall), F1-점수(score)는 식 (2)~(5)와 같은 방법으로 계산된다.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$F1-score = \frac{2 \times Precision \times Recall}{\mathrm{P}recision + Recall} \quad (5)$$

여기서, TP는 True positive로, 모델이 양성(공격)으로 올바르게 예측한 데이터 개수를, TN은 True negative로, 모델이 음성(정상)으로 올바르게 예측한 데이터 개수를, FP는 False positive로 실제 양성 클래스를 음성으로 잘못 예측한 데이터 개수를, 마지막으로 FN은 실제 음성 클래스를 양성으로 잘못 예측한 데이터 개수를 각 의미한다. 이때, 정확도는 모든 데이터 중, 모델이 올바르게 분류한 데이터의 비율을 나타내, 정밀도는 양성으로 예측한 데이터 중, 실제 양성의 비율을 나타낸다. 재현율은 실제 양성 중, 모델이 양성으로 예측한 비율을 나타내며, F1-점수는 정밀도와 재현율의 조화 평균을 계산한 값이다. 본 연구에서는 다중 클래스 분류에서의 정확도, 정밀도, 재현율, F1-점수를 계산하기 위하여 모든 클래스를 이진 클래스 분류라고 가정하여, 각 클래스에 대하여 상기의 수식을 통해 구한 점수들의 평균을 사용한다.

표 1. 공격 유형 분류 정확도 비교
Table 1. Comparison of attack category classification accuracy

No.	Model	Accurucy(%)
#1	Logistic regression	77.63
#2	SVM	80.97
#3	1D CNN-based method	92.37
#4	Transfer learning-based method	**94.21**

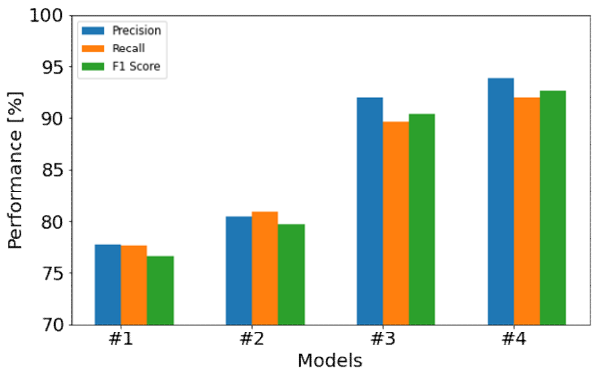

그림 9. 모델에 따른 네트워크 침입 탐지 성능 비교
Fig. 9. Comparison of network intrusion detection performance by model

식 (2)~(5)의 성능 지표들을 활용한 표 1과 그림 9에서의 실험 결과를 통해 제안된 네트워크 침입 탐지 모델이 가장 우수한 성능을 보이는 것을 실험적으로 확인할 수 있다.

Ⅵ. 결론 및 향후 과제

본 논문에서는 네트워크 상에서 발생할 수 있는 다양한 종류의 사이버 공격을 탐지하기 위해, 전이 학습 기법을 기반으로 네트워크 침입 탐지 모델을 제안하였다. 실험 결과, 전이 학습 기법을 적용한 네트워크 침입 탐지 모델에서 약 94.21%의 정확도 성능을 얻을 수 있었으며, 전이 학습 기법을 적용하지 않은 모델 보다 약 1.84% 더 우수한 성능을 보이는 것을 확인할 수 있었다. 이는 네트워크 침입 탐지 분야에서 전이 학습을 적용하는 것이 새로운 유형의 공격을 더욱 효과적으로 학습하여, 정확하게 대응할 수 있음을 입증한다.

향후에는 공격 유형 데이터에서 발생하는 클래스 불균형에 따른 오탐율을 감소하기 위하여, 모델 학습 시 클래스 불균형에 따른 가중치를 적용한 전이 학습 기반의 네트워크 침입 탐지 모델에 대한 연구를 수행할 예정이다.

References

[1] I. Alrashdi, A. Alqazzaz, E. Aloufi, R. Alharti, M. Zohdy, and H. Ming, "AD-IoT: Anomaly Detection of IoT Cyber attacks Smart City Using Machine Learning", 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, pp. 305-310, Jan. 2019. https://doi.org/10.1109/CCWC.2019.8666450.

[2] T. H. Kim and S. H. Kim, "An Intrusion Detection System based on the Artificial Neural Network for Real Time Detection", Convergence Security Journal, Vol. 17, No. 1, pp. 31-38, Mar. 2017.

[3] J. H. Lee, J. G. Pak and M. S. Lee, "Network

Intrusion Detection System Using Feature Extraction Based on AutoEncoder in IoT environment", 2020 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea, pp. 1282-1287, Oct. 2020. https://doi.org/10.1109/ICTC49870.2020.9289253.

[4] B. J. Min, J. H. Yoo, S. S. Kim, D. I. Shin, and D. K. Shin, "Network Intrusion Detection with One Class Anomaly Detection Model based on Auto Encoder", Journal of Internet Computing and Services, Vol. 22, Vo. 1, pp. 13-22, Feb. 2021. https://doi.org/10.7472/jksii.2021.22.1.13.

[5] K. Shiomoto, "Network Intrusion Detection System Based on an Adversarial Auto-Encoder with Few Labeled Training Samples", Journal of Network and Systems Management, Vol. 31, No. 5, Sep. 2022. https://doi.org/10.1007/s10922-022-09698-w.

[6] T. Hirakawa, K. Ogura, B. B. Bista, and T. Takata, "A Defense Method against Distributed Slow HTTP DoS Attack", 2016 19th International Conference on Network-Based Information Systems (NBiS), Ostrava, Czech Republic, pp. 152-158, Sep. 2016. https://doi.org/10.1109/NBiS.2016.58.

[7] J. H. Jafarian, E. Al-Shaer, and Q. Duan, "An Effective Address Mutation Approach for Disrupting Reconnaissance Attacks", IEEE Transactions on Information Forensics and Security, Vol. 10, No. 12, pp. 2562-2577, Dec. 2015. https://doi.org/10.1109/TIFS.2015.2467358.

[8] A. Tekerek, C. Gemci, and O. F. Bay, "Development of a Hybrid Web Application Firewall to Prevent Web Based Attacks", 2014 IEEE 8th International Conference on Application of Information and Communication Technologies (AICT), Astana, Kazakhstan, pp. 1-4, Oct. 2014. https://doi.org/10.1109/ICAICT.2014.7035910.

[9] S. W. Lim and M. S. Yoo, "Comparison on Machince Learning Techniques for Intrusion Detection in Wireless Sensor Network", J-KICS, Vol. 47, No. 11, pp. 1804-1814, Nov. 2022. https://doi.org/10.7840/kics.2022.47.11.1804.

[10] L. Ashiku and C. Dagli, "Network Intrusion Detection System using Deep Learning", Procedia Computer Science, Vol. 185, pp. 239-247, Jun. 2021. https://doi.org/10.1016/j.procs.2021.05.025.

[11] K. H. Kang, "Network Anomaly Detection Technologies Using Unsupervised Learning Auto Encoders", Journal of the Korea Institute of Information Security & Cryptology, Vol. 30, No. 4, pp. 617-629, Aug. 2020. https://doi.org/10.13089/JKIISC.2020.30.4.617.

[12] N. Koroniotis, N. Moustafa, E, Sitnikova, and B. Turnbull, "Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset", Future Generation Computers Systems, Vol. 100, pp. 779-796, Nov. 2019. https://doi.org/10.1016/j.future.2019.05.041.

[13] N. Moustafa and J. Slay, "UNSW-NB15: A Comprehensive Data set for Network Intrusion Detection systems", 2015 Military Communications and Information Systems Conference (MilCIS), Canberra, ACT, Australia, pp. 1-6, Nov. 2015. https://doi.org/10.1109/MilCIS.2015.7348942.

[14] M. S. Kim, H. J. Yang, S. H. Kim, and W. P. Cheah, "Improved Focused Sampling for Class Imbalance Problem", Korea Information Processing Society. Part B, Vol. 14B, No. 4, pp. 287-294, Aug. 2007. https://doi.org/10.3745/KIPSTB.2007.14-B.4.287.

[15] J. Stefanowski, "Overlapping, Rare Examples and Class Decomposition in Learning Classifiers from Imbalanced Data", Institute of Computing Science, Vol. 13, pp. 277-306, Aug. 2013. https://doi.org/10.1007/978-3-642-28699-5_11.

[16] P. Wu, H. Guo, and R. Buckland, "A Transfer Learning Approach for Network Intrusion Detection", 2019 IEEE 4th International Conference on Big Data Analytics (ICBDA),

Suzhou, China, pp. 281-285, Mar. 2019. https://doi.org/10.1109/ICBDA.2019.8713213.

[17] C. H. Park, J. H. Lee, Y. S. Kim, J. G. Park, H. J. Kim, and D. W. Hong, "An Enhanced AI-Based Network Intrusion Detection System Using Generative Adversarial Networks", IEEE Internet of Things Journal, Vol. 10, No. 3, pp. 2330-2345, Feb. 2023. https://doi.org/10.1109/JIOT.2022.3211346.

[18] Y. Fan, Y. Li, M. Zhan, H. Cui, and Y. Zhang, "IoTDefender: A Federated Transfer Learning Intrusion Detection Framework for 5G IoT", 2020 IEEE 14th International Conference on Big Data Science and Engineering (BigDataSE), Guangzhou, China, pp. 88-95, Dec. 2020. https://doi.org/10.1109/BigDataSE50710.2020.00020.

[19] A. Singla, E. Bertino, and D. Verma, "Overcoming the Lack of Labeled Data: Training Intrusion Detection Models Using Transfer Learning", 2019 IEEE International Conference on Smart Computing (SMARTCOMP), Washington, DC, USA, pp. 69-74, Jun. 2019. https://doi.org/10.1109/SMARTCOMP.2019.00031.

[20] J. Y. Hwang, B. A. Choi, J. H. Lee, and J. H. Bae, "A Study on Transfer Learning-based Ensemble Model for Classification of Maneuver Weapon System", Journal of KIIT, Vol. 19, No. 12, pp. 1-10, Dec. 2021. http://dx.doi.org/10.14801/jkiit.2021.19.12.1.

[21] Y. H. Byeon and K. C. Kwak, "A Transfer Learning and Performance Comparison of Deep Learning Models for Pedestrian Classification under Automobile Driving Environment", Journal of KIIT, Vol. 16, No. 10, pp. 83-92, Oct. 2018. http://dx.doi.org/10.14801/jkiit.2018.16.10.83.

[22] K. O`Shea and R. Nash, "An Introduction to Convolutional Neural Networks", arXiv preprint arXiv:1511.08458v2, pp. 1-11, Nov. 2015. https://doi.org/10.48550/arXiv.1511.08458.

[23] A. F. Agarap, "Deep Learning using Rectified Linear Units (ReLU)", arXiv preprint arXiv:1803.08375v2, pp. 1-7, Feb. 2019. https://doi.org/10.48550/arXiv.1803.08375.

[24] L. F. S. Scabini and O. M. Bruno, "Structure and performance of fully connected neural networks: Emerging complex network", arXiv preprint arXiv:2107.14062v1, pp. 1-18, Jul. 2021. https://doi.org/10.48550/arXiv.2107.14062.

[25] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization", arXiv preprint arXiv:1412.6980v9, pp. 1-15, Jan. 2017. https://doi.org/10.48550/arXiv.1412.6980.

[26] C. Darken, J. Chang, and J. Moody, "Learning rate schedules for faster stochastic gradient search", Neural Networks for Signal Processing II Proceedings of the 1992 IEEE Workshop, Helsingoer, Denmark, pp. 3-12, Aug. 1992. https://doi.org/10.1109/NNSP.1992.253713.

저자소개

주 승 세 (SeungSae Joo)

2021년 3월 ~ 현재 :
대구가톨릭대학교
AI빅데이터공학과 학사과정
관심분야 : 인공지능,
딥러닝/머신러닝

이 성 혁 (SungHyuk Lee)

2021년 3월 ~ 현재 :
대구가톨릭대학교
AI빅데이터공학과 학사과정
관심분야 : 인공지능,
딥러닝/머신러닝

유 재 학 (Jaehak Yu)

2001년 2월 : 건국대학교
전산과학과(학사)
2003년 2월 : 고려대학교
전산학과(이학석사)
2010년 2월 : 고려대학교
전산학과(이학박사)
2010년 10월 ~ 현재 :
한국전자통신연구원 책임연구원
관심분야 : 데이터마이닝, 인공지능, 네트워크 암호 트래픽 분석, 네트워크 보안, 기계학습

문 대 성 (Daesung Moon)

2007년 2월: 고려대학교
전산학과(공학박사)
2000년 12월 ~ 현재 :
한국전자통신연구원 책임연구원
2009년 3월 ~ 현재 :
과학기술연합대학원대학교(UST)
정보통신공학 전공책임교수
관심분야 : 정보보호, 네트워크보안, AI 보안, 인공지능, 딥러닝

배 지 훈 (Ji-Hoon Bae)

2000년 2월 : 경북대학교
전자·전기공학부(공학사)
2002년 2월 : 포항공과대학교
전자컴퓨터공학부(공학석사)
2016년 2월 : 포항공과대학교
전자·전기공학과(공학박사)
2002년 1월 ~ 2019년 8월 :
한국전자통신연구원 책임연구원
2019년 9월 ~ 현재 : 대구가톨릭대학교 AI빅데이터공학과 조교수
2021년 3월 ~ 현재 : 대구가톨릭대학교 SW중심대학사업단 SW기초교육센터장
관심분야 : 인공지능, 딥러닝/머신러닝, 레이다 영상 및 신호처리, 최적화 기법