J. Internet Comput. Serv.
ISSN 1598-0170 (Print) / ISSN 2287-1136 (Online)
http://www.jics.or.kr

Extra Tree와 ANN을 활용한 이상 탐지 및 공격 유형 분류 메커니즘☆

Anomaly detection and attack type classification mechanism using Extra Tree and ANN

김 민 규[1] 한 명 묵[2*]
Min-Gyu Kim Myung-Mook Han

요 약

이상 탐지는 일반적인 사용자들의 데이터 집합 속에서 비정상적인 데이터 흐름을 파악하여 미리 차단하는 방법이다. 기존에 알려진 방식은 이미 알려진 공격의 시그니처를 활용하여 시그니처 기반으로 공격을 탐지 및 방어하는 방식인데, 이는 오탐율이 낮다는 장점이 있지만 제로 데이 취약점 공격이나 변형된 공격에 대해서는 매우 취약하다는 점이 문제점이다. 하지만 이상 탐지의 경우엔 오탐율이 높다는 단점이 존재하지만 제로 데이 취약점 공격이나 변형된 공격에 대해서도 식별하여 탐지 및 차단할 수 있다는 장점이 있어 관련 연구들이 활발해지고 있는 중이다. 본 연구에서는 이 중 이상 탐지 메커니즘에 대해 다뤘다. 앞서 말한 단점인 높은 오탐율을 보완하며 그와 더불어 이상 탐지와 분류를 동시에 수행하는 새로운 메커니즘을 제안한다. 본 연구에서는 여러 알고리즘의 특성을 고려하여 5가지의 구성으로 실험을 진행하였다. 그 결과로 가장 우수한 정확도를 보이는 모델을 본 연구의 결과로 제안하였다. Extra Tree와 Three layer ANN을 동시에 적용하여 공격 여부를 탐지한 후 공격을 분류된 데이터에 대해서는 Extra Tree를 활용하여 공격 유형을 분류하게 된다. 본 연구에서는 NSL-KDD 데이터 세트에 대해서 검증을 진행하였으며, Accuracy는 Normal, Dos, Probe, U2R, R2L에 대하여 각각 99.8%, 99.1%, 98.9%, 98.7%, 97.9%의 결과를 보였다. 본 구성은 다른 모델에 비해 우수한 성능을 보였다.

☞ 주제어 : 익스트림 랜덤 포레스트, 인공 신경망, 이상 탐지, 이상 탐지 및 공격 유형 분류, 네트워크 침입 탐지

ABSTRACT

Anomaly detection is a method to detect and block abnormal data flows in general users' data sets. The previously known method is a method of detecting and defending an attack based on a signature using the signature of an already known attack. This has the advantage of a low false positive rate, but the problem is that it is very vulnerable to a zero-day vulnerability attack or a modified attack. However, in the case of anomaly detection, there is a disadvantage that the false positive rate is high, but it has the advantage of being able to identify, detect, and block zero-day vulnerability attacks or modified attacks, so related studies are being actively conducted. In this study, we want to deal with these anomaly detection mechanisms, and we propose a new mechanism that performs both anomaly detection and classification while supplementing the high false positive rate mentioned above. In this study, the experiment was conducted with five configurations considering the characteristics of various algorithms. As a result, the model showing the best accuracy was proposed as the result of this study. After detecting an attack by applying the Extra Tree and Three-layer ANN at the same time, the attack type is classified using the Extra Tree for the classified attack data. In this study, verification was performed on the NSL-KDD data set, and the accuracy was 99.8%, 99.1%, 98.9%, 98.7%, and 97.9% for Normal, Dos, Probe, U2R, and R2L, respectively. This configuration showed superior performance compared to other models.

☞ keyword : Extreme Random Forest, Artificial Neural Network, Anomaly Detection, Anomaly Detection and Attack type Classification, Network Intrusion Detection

1. 서 론

대부분의 산업에서 디지털화가 진행됨에 따라 네트워

1 Department of Computer Engineering, Gachon University, Seongnam-si, 13120, Korea.
2 Department of Software, Gachon University, Seongnam-si, 13120, Korea.
* Corresponding author (mmhan@gachon.ac.kr)
[Received 8 August 2022, Reviewed 13 August 2022(R2 12 September 2022), Accepted 12 October 2022]

☆ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음 (IITP-2022-2020-0-01602)

크 데이터는 급속도로 증가했다. 한 보안관제센터에서 발생하는 보안 관제 이벤트는 초당 15만 건, 하루에 10억 건씩 발생하고 있다[1]. 이처럼 하루가 다르게 생성되는 데이터의 수가 폭증하고 있는 상황에서 유무선 네트워크 데이터에서 비정상적인 데이터를 식별하고 탐지하는 것은 매우 어려운 일이다. 이런 상황에서 자연스레 네트워크 보안에 대한 중요도는 높아지며 관련 연구가 활성화되고 있는 추세이다. 네트워크 공격은 인가되지 않은 사용자가 접근하여 정상적인 네트워크 활동을 방해하여 수행 능력을 약화시키거나 불가능하게 하거나 네트워크 활동을 분석하여 원하는 정보를 얻는 일련의 과정을 의미한다. 최근 네트워크의 규모가 확대되며 더 많은 취약점이 존재하게 되었다. 이를 사전에 차단하기 위한 네트워크 침입 탐지 기술은 네트워크 내의 트래픽을 분석하여 비정상적인 트래픽을 식별하여 사전에 차단하는 기술이다. 일반적으로 서비스 거부 공격, 싱크 플루딩, 핑, 스머프 공격 등을 식별하여 차단할 수 있다. 공격자들은 네트워크 공격을 통해 개인정보나 암호화폐, 자산, 기밀 정보 등의 가치 있는 정보나 자산을 탈취한다.

전통적으로는 시그니처를 이용한 탐지 방법이 주를 이뤘었는데 이는 사전에 알려진 유형에 대해서만 공격식별이 가능하다[2]. 그렇기 때문에 오탐률은 낮지만, 기존에 없던 유형의 공격에 대해서는 사전에 차단하기 힘들다는 것이 큰 취약점이다. 그에 반해 이상 탐지를 활용한 침입탐지 시스템은 사전에 알려지지 않은 유형을 식별할 수 있다는 점은 장점이지만 정상적으로 작동하는 데이터를 비정상적인 데이터로 식별할 수도 있다는 단점이 있다[3]. 이를 해결하기 위해서 활발한 연구가 진행되고 있다. 하지만 연구하기에 적합한 데이터가 부족하다는 점과 데이터의 불균형 등으로 인해서 좋은 결과를 내지 못하는 경우가 많다.

그에 따라 본 연구에서는 NSL-KDD 데이터를 활용하여 네트워크 내에서의 이상 탐지를 식별하고 분류하는 2단계 이상 탐지 메커니즘을 제안한다. 본 연구에서 5가지의 다양한 알고리즘을 조합을 통해 실험하여 우수한 성능을 검증하는 과정을 보였다.

결과로는 Extra Tree와 ANN을 조합하여 좋은 성능을 내는 모델을 제안했다. 그 결과 Accuracy는 Normal, Dos, Probe, U2R, R2L에 대하여 각각 99.8%, 99.1%, 98.9%, 98.7%, 97.9%의 결과를 도출했다. 본 조합의 경우 다른 조합에 비해 우수한 결과를 보였다.

본 논문의 구성은 다음과 같다. 2장에서는 침입 탐지 및 분류에 대한 설명 및 관련 연구에 대한 소개를 하고, 3장에서는 제안하는 모델에 대해 구체적으로 설명한다. 4장에서는 해당 모델의 실험에 대해 소개하며 데이터 세트 및 실험 결과를 분석한다. 마지막으로 5장에서는 본 논문의 결론을 도출하며 향후 연구에 대해 언급하며 끝을 맺는다.

2. 침입 탐지 및 공격 유형 분류 관련 연구

2.1 침입 탐지 및 공격 유형 분류

침입 탐지 시스템은 컴퓨터 시스템 자원에서의 보안 요소를 저해하는 모든 행위를 탐지하고 식별하는 시스템을 의미한다. 이는 시스템의 보안을 위해 위협을 탐지하거나 데이터의 유출이 발생했을 때 경로를 추적하는 등의 역할을 하는데 최근 해당 연구에 인공지능과 기계학습의 적용이 많아지고 있다. 이상 탐지에 있어 NIDS (Network Intrusion Detection System)에 ANN 기술의 적용으로 이상 탐지의 성능 향상에 크게 기여하고 있다. 이 NIDS는 네트워크에서의 지정된 지점을 지나가는 패킷들을 모두 분석이 가능한지, 그리고 오탐 없이 정확한 탐지가 가능한지가 확실하지 않다는 점이 문제점으로 대두되고 있다.

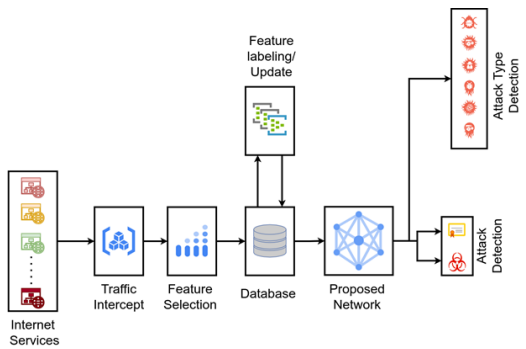
2.2 관련 연구

Yisroel et al.[4] 연구팀은 효율적으로 Plug & Play 할 수 있게 설계된 신경망 기반의 NIDS를 제안했다. 이는 모든 네트워크 채널의 동작을 효율적으로 추적하고 이상 탐지를 위해 Auto Encoder의 앙상블을 사용하여 해당 작업을 수행한다. 해당 연구에서는 온라인 머신러닝 프로세스에 대해서 자세히 논의 했으며 이를 탐지 및 런타임 성능 측면에서 평가한 결과 기존의 다른 알고리즘과 비슷하거나 경우에 따라 더 우수한 결과를 도출한다. 그뿐만 아니라 해당 알고리즘은 Raspberry PI의 단일 코어에서 실행하기에 충분히 효율적이며 더 성능이 좋은 CPU나 GPU에서는 훨씬 더 좋은 성능을 보일 수 있다.

Sunwoo. Ahn et al.[5] 연구팀은 ANN 알고리즘의 가중치 양자화를 제안하여 복잡한 연산을 하는데 필요한 메모리를 줄였다. 결과는 연구 이전에 비해 최대 약 30배의 속도 향상을 가져왔으며, Raspberry PI에서 프로토타입을 구현하여 평가했다. 그뿐만 아니라 NIDS용 최신 프로세서에 SIMD(Single Instruction Multiple Data) 엔진을 적용하여 기존에 비해 약 66배의 속도 향상을 이뤄냈으며 연구 성과를 입증했다.

3. 제안 모델

본 연구는 네트워크상에서 비정상적인 데이터를 탐지하고 분류하기 위한 방법에 대한 연구를 수행하였다. 제안하는 모델의 프레임워크는 아래 그림 1과 같이 4단계로 구성하였다.



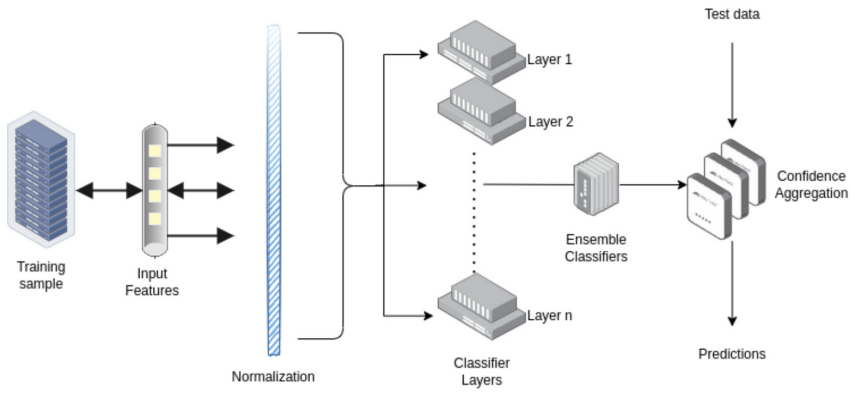
(그림 1) 제안 모델 구조
(Figure 1) Proposed Model Architecture

우선 Sniffing 도구를 활용하여 네트워크 트래픽을 얻어 Wireshark나 Tcpdump 등으로 캡처한 패킷을 이용해 얻은 특징을 데이터 세트에 저장한다. 추출한 특징은 전처리 과정을 거쳐서 중복을 제거하고, 다양한 특징을 정규화하여 Ont-Hot Encoding을 한다.

(표 1) 패킷별 공격 유형 분류
(Table 1) Categorize attack types by packet

종류	유형
DOS	Back, Land, Neptune, Pod, Smurf, Teardrop, Mailbomb, Processtable, Udpstorm, Apache2, Worm
Probe	Satan, IPsweep, Nmap, Portsweep, Mscan, Saint
R2L	Guess_password, Ftp_write, Imap, Phf, Multihop, Warezmaster, Xlock, Xsnoop, Snmpguess, Snmpgetattack, Httptunnel, Sendmail, Named
U2R	Buffer_overflow, Loadmodule, Rootkit, Perl, Sqlattack, Xterm, Ps

데이터 세트에서의 유형은 다음과 같다. 정상적인 트래픽은 Normal로 분류하고, DOS는 일반 이용자들이 웹 서비스에 접근하지 못하게 막는 행위인데 이는 네트워크나 서버를 플러딩하는 방식으로 웹 서비스를 충돌시키거나 리소스를 고갈시켜 서비스를 정지시키는 방식을 통해 공격하는 방법으로, Back, Land, Neptune, Pod, Smurf, Teardrop, Mailbomb, Processtable, Udpstorm, Apache2, Worm 유형을 포함한다. Probe는 실제로 공격하기에 앞서 시스템 내부의 패킷 등과 같은 다양한 자료를 수집하는 과정으로, Satan, IPsweep, Nmap, Portsweep, Mscan, Saint를 포함한다. R2L은 Remote to Local의 약자로 시스템에 아무런 권한이 없는 이용자가 시스템의 외부에서 강제로 접근 권한을 얻기 위한 행위로, Guess_password, Ftp_write, Imap, Phf, Multihop, Warezmaster, Xlock, Xsnoop, Snmpguess, Snmpgetattack, Httptunnel, Sendmail, Named 패킷을 포함한다. U2R은 User to Root의 약자로 시스템 사용자가 강



(그림 2) 제안 네트워크
(Figure 2) Proposed Network

제로 Root 권한을 취득하기 위한 공격으로 Buffer_overflow, Loadmodule, Rootkit, Perl, Sqlattack, Xterm, Ps를 모두 포함한다.

위 그림 2는 제안 모델에 적용되는 제안 Network 프레임워크이다.

분류기의 앙상블을 통해 모델은 공격 여부를 식별하며 식별된 공격의 유형을 분류하게 된다. 머신러닝 방식은 일반적으로 규칙을 기반으로 구성하는 경우가 대부분이기에 네트워크상에서 Instance 속성을 잘못 지정할 수 있다는 문제점이 있다. 그 불확실성을 해결하기 위해 분류기를 통합하여 함께 사용하는 방법을 제안한다.

본 연구에서 Extra Tree 알고리즘의 경우엔 다른 알고리즘에 비해 런타임이 매우 짧다는 특징이 있음을 확인했으며 아래 표 2에서 각 알고리즘들의 장단점을 확인할 수 있다. 따라서 런타임 측면에서의 장점을 극대화하기 위해 모델 후보군에 두고 다양한 조합을 찾았다. 앙상블에 사용하는 분류기 중 하나는 Three Layer ANN을 활용하며, 그 이유는 ANN 알고리즘의 경우 평균 재현율의 결과가 우수하기 때문이다. 그리고 ML 방식은 일반적으로 규칙을 기반으로 구성하는 경우가 대부분이기에 네트워크상에서 인스턴스 속성을 잘못 지정할 수 있다는 문제점이 있고, 그 불확실성을 해결하기 위해 ANN 알고리즘을 분류기 조합에 포함하였다.

(표 2) 알고리즘별 장단점
(Table 2) Advantages and disadvantages of each algorithm

알고리즘	장점	단점
Extra Tree	- Random Forest에 비해 낮은 Bias - 낮은 Variance - 빠른 연산속도	- 무작위성이 더 커졌기에 더 많은 트리를 훈련해야 함
ANN	- 변수가 많거나 입출력 변수 사이의 복잡한 비선형 관계에 유용 - 우수한 평균 재현율	- 최적의 파라미터 찾기 어려움 - 긴 학습시간 - Overfitting 가능
LSTM	- 순서가 있는 데이터에 적용하기 적합	- 간단한 문제에 적용하기 부적합
SVM	- 분류 및 예측에 동시에 사용 가능 - Overfitting이 덜함 - 높은 정확도	- 모델 구축 시간이 오래 걸림 - 결과에 대한 설명성 부족
Regression	- 빠른 학습 속도 - 빠른 예측 속도 - 대용량 데이터 처리	- 계수 값의 분석이 어려움
MLP	- 복잡한 모델도 처리 가능 - 우수한 성능	- 학습 시간이 오래 걸림 - 데이터 전처리의 영향을 크게 받음

여러 조합에서 41가지의 특징을 활용하여 공격을 식별해본 후, 특정 15가지의 특징만으로 공격을 식별하는 실험을 추가적으로 진행한 결과 특정 특징만을 활용하여 공격을 식별 및 분류한 결과가 비교적 우수한 성능을 보였다. 본 연구에서는 Extra Tree + ANN, LSTM, SVM + SVM, Regression + ANN, MLP 총 5가지의 알고리즘 조합을 활용하여 결과를 비교하는 실험을 진행했다.

4. 실 험

4.1 실험 환경

본 연구를 수행함에 있어 다음과 같은 실험 환경으로 실험을 수행했다. 실험 환경의 운영체제는 Windows 10 Pro였고, 자세한 실험 환경은 Intel(R) Core(TM) i7-8700 CPU @ 3.20GHz의 CPU, NVIDIA GeForce GTX 1050 Ti 의 GPU, 16GB RAM 환경 등에서 수행했다.

4.2 NSL-KDD Dataset

NSL-KDD는 기존의 KDD 99 데이터 세트의 문제점을 보완하여 생성된 데이터 세트이다[6, 7]. 해당 데이터 세트는 여전히 다양한 문제점이 존재하지만 KDD 99에 비해 중복되는 레코드의 제거로 데이터의 편차를 줄였으며, 현재 다양한 침입 탐지 연구들에 적용되고 있는 추세이다. 위 데이터 세트는 Train 및 Test 데이터의 비율이 적절하게 분배되어 있으며 Normal 데이터와 더불어 DOS, Probe, U2R, R2L에 대한 공격 데이터들을 포함하고 있고, Duration, Protocol Type, Service 등을 포함하여 총 41개의 특성을 가지고 있다.

해당 데이터 세트에서 특성 열에 존재하는 공격 라벨과 공격 유형을 따로 분리해서 훈련에 정답 라벨로 활용하여 훈련했다. 또한 머신러닝 알고리즘에 데이터를 주입하기 위해서 열 데이터 유형을 정수 혹은 실수 데이터로 변환한 후 One-Hot Encoding을 통해 전처리했다.

4.3 실험 결과

해당 연구에서는 앞서 선별한 알고리즘 조합 5가지를 활용하여 실험을 진행하였으며 최상의 결과를 낼 수 있는 조합을 실험 결과로 제공하고자 한다.

아래 표 3는 실험 결과이며, 공격 탐지와 공격 유형을 분류하기 위한 모델을 조합한 실험이다.

본 논문의 제안 모델은 Extra Tree와 ANN을 융합한 모델로 그림 2와 같은 구성을 보인다. 이는 전반적으로 우수한 성능을 보였으며 모든 특성을 활용하여 탐지를 진행하는 것보다 일부 특성들을 활용하여 얻은 결과가 더 좋은 결과를 보였다. 따라서 본 연구에서 해당 조합을 제안하였으며, 앞서 설명한 것처럼 그림 2와 같은 구성을 보인다.

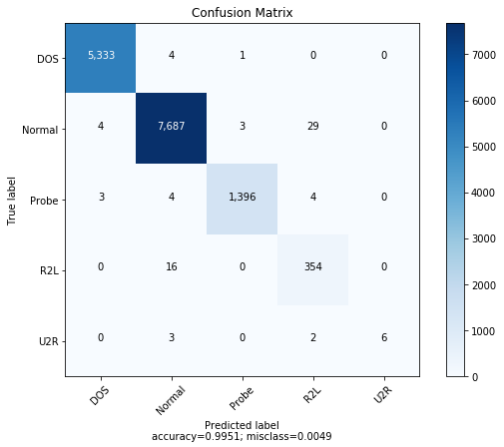
(표 3) 각 조합별 실험 결과
(Table 3) Experimental results for each combination

조합	클래스명	정확도 (41개 특징)	정확도 (15개 특징)
Extra Tree + ANN	Normal	99.1	99.8
	DOS	98.7	99.1
	Probe	97.6	98.9
	U2R	97.5	98.7
	R2L	96.8	97.9
LSTM	Normal	97.9	98.5
	DOS	82.4	88.3
	Probe	80.2	86.0
	U2R	73.9	75.5
	R2L	87.6	88.9
SVM+SVM	Normal	98.1	98.9
	DOS	97.8	98.6
	Probe	90.7	91.3
	U2R	93.7	95.9
	R2L	91.8	93.9
Regression+ ANN	Normal	88.9	91.9
	DOS	82.7	89.5
	Probe	82.1	85.4
	U2R	73.1	80.7
	R2L	80.8	89.0
MLP	Normal	80.3	85.9
	DOS	72.7	75.0
	Probe	70.9	75.1
	U2R	70.7	74.3
	R2L	69.8	71.1

실험 결과는 41가지 특성을 활용한 경우와 15가지 특성을 활용한 경우의 정확도로 평가를 진행했다.

MLP 알고리즘의 경우 모델링의 완성도에 따라 성능의 차이가 발생하기에 우수한 정확도를 기대하였으나, 타 알고리즘 조합에 비해 좋지 못한 결과를 도출하였다. 대용량 데이터 처리 및 많은 변수를 활용한 빠른 학습을 위하여 조합한 Regression과 ANN의 조합은 예상보다 우수한 정확도를 확인할 수 없었다. 분류 및 예측을 동시에 적용하기에 적합한 SVM의 경우엔 예상한대로 높은 정확도를 보였으나 모델 구축에 시간이 오래 걸린다는 단점이 존재하였다. 시계열 데이터에 다양하게 적용 가능한 LSTM 알고리즘은 예상보다 높은 정확도를 확인할 수 있었다.

Extra Tree 알고리즘과 ANN의 조합은 가장 우수한 결과를 보였기에 본 조합을 제안한다. Extra Tree 알고리즘은 다른 방식들에 비해 런타임 측면에서의 우수함이 있고, ANN 알고리즘을 활용하여 재현율 측면에서의 이점이 있어 다른 알고리즘에 비해 적합한 조합이다. 그뿐만 아니라 정확도 측면에서도 가장 좋은 결과를 보였기에 본 연구의 결과는 검증되었다. 본 연구에서는 NSL-KDD를 통해서 제안하는 모델의 성능을 검증하였으며, 그 결과는 그림 3의 Confusion Matrix 결과를 확인할 수 있다. 또한, ANN 알고리즘의 경우 Epoch 값을 100으로 지정하여 Overfitting을 피할 수 있게 하였고, 신경망 데이터를 훈련하는데 일반적으로 가장 많이 쓰이는 Cross Entropy 방법을 Loss 함수로 적용하였다[8]. Learning rate는 0.1로 지정하여 모델을 학습했다.



(그림 3) 제안 네트워크의 Confusion Matrix 결과
(Figure 3) Confusion Matrix Results of the Proposal Network

5. 결론 및 향후 연구

일반적으로 이상 탐지 모델들은 단일 알고리즘으로 구성되어 네트워크 대역 망에서 비정상적인 데이터를 식별하거나 공격 유형을 분류하는 방법으로 사용된다. 기존의 방식에서는 새로운 형식의 알고리즘이 개발되지 않는 한 성능 개선에 많은 제약사항이 존재한다. 본 연구에

서는 이런 문제를 해결하고자 네트워크 이상 탐지를 새롭게 융합한 방식을 제안하였다.

첫 번째는 효과적인 이상 탐지를 위하여 여러 알고리즘을 비교하여 가장 우수한 성능을 보인 조합을 찾았다. 총 5가지의 조합 중 Extra Tree + ANN을 통해 가장 우수한 성능을 확인할 수 있었다. 제안된 모델은 앞서 소개한 NSL-KDD 데이터 세트에 대하여 검증되었다.

하지만 본 연구의 결과는 해당 데이터 세트에 대해서만 결과를 확인하였다. 추후에 다른 네트워크 데이터에 대해서 성능을 추가적으로 검증할 필요가 있으며, 향후 연구로는 본 연구보다 다양한 알고리즘 조합을 활용하고, 균형감 있는 데이터 세트를 구축하기 위한 GAN이나 Auto Encoder를 활용한 연구도 수행할 것이다[9, 10]. 다음과 같은 데이터를 활용한다면 더욱 향상된 탐지 및 분류 성능을 가져올 것으로 기대된다. 따라서 해당 연구는 추후 부족한 부분을 보완하여 연구를 진행할 예정이다.

본 연구에서는 기존 연구들과는 다르게 이상 탐지로 그치지 않고 공격 유형을 분류했음에 의의가 있다. 관련 연구들은 Raspberry PI라는 환경에서 경량화가 필요한 상황에서 우수한 성능을 보였으나, 본 연구는 경량화되지 않은 일반적인 PC에서 적용이 가능하다는 점에서 제약의 차이도 존재했다. 추후 연구로 본 연구의 모델을 경량화에도 도전하고자 한다.

참고문헌(Reference)

[1] K. Kug, B. Gong, "Security technology development trend using artificial intelligence", Institute of Information and Communication Planning and Evaluation Weekly Technology Trend, pp. 2-15, 2019.
https://www.iitp.kr/kr/1/knowledge/periodicalViewA.it?searClassCode=B_ITA_01&masterCode=publication&identifier=1095

[2] G. Creech and J. Hu, "A Semantic Approach to Host-Based Intrusion Detection Systems Using Contiguousand Discontiguous System Call Patterns", IEEE Transactions on Computers, vol. 63, no. 4, pp. 807-819, 2014.
https://doi.org/10.1109/TC.2013.13

[3] N. Moustafa and, J. Slay, "A hybrid feature selection for network intrusion detection systems: Central points", 16th Australian Information Warfare Conference, pp. 5-13, 2015.
http://dx.doi.org/10.13140/RG.2.1.3905.5122

[4] Y. Mirsky, T. Doitshman, Y. Elovici and A. Shabtai, "Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection", Network and Distributed Systems Security Symposium(NDSS), 2018.
https://doi.org/10.48550/arXiv.1802.09089

[5] S. Ahn, H. Yi, Y. Lee, W. R. Ha, G. Kim and Y. Paek, "Hawkware: Network Intrusion Detection based on Behavior Analysis with ANNs on an IoT Device" 57th ACM/IEEE Design Automation Conference (DAC), pp. 1-6, 2020.
https://doi.org/10.1109/DAC18072.2020.9218559

[6] M. Tavallaee, E. Bagheri, W. Lu and A. Ghorbani, "A Detailed Analysis of the KDD CUP 99 Data Set", 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, pp. 1-6, 2009.
https://doi.org/10.1109/CISDA.2009.5356528

[7] M. Tavallaee, E. Bagheri, W. Lu and A. Ghorbani, "Nsl-kdd dataset", 2012.
http://www.unb.ca/research/iscx/dataset/iscx-NSL-KDD-dataset.html

[8] X. Li, D. Chang, T. Tian and J. Cao, "Large-Margin Regularized Softmax Cross-Entropy Loss.", IEEE Access, vol. 7, pp. 19572-19578, 2019.
https://doi.org/10.1109/ACCESS.2019.2897692

[9] A. Liu, Y. Wang and T. Li, "SFE-GACN: A novel unknown attack detection under insufficient data via intra categories generation in embedding space", Computers & Security, vol. 105, 2021.
https://doi.org/10.48550/arXiv.2004.05693

[10] Y. Kim, "Self-supervised auto-encoder for anomaly detection", Master's diss, Pohang University of Science and Technology, 2019.2.
http://www.riss.kr/link?id=T15273279

◐ 저 자 소 개 ◑

김 민 규(Min-Gyu Kim)
2021년 가천대학교 컴퓨터공학과(공학사)
2022년 가천대학교 대학원 IT융합공학과(공학석사)
관심분야 : 정보보호, 모의해킹, 취약점분석, 보안컨설팅
E-mail : alsq0506@gachon.ac.kr

한 명 묵(Myung-Mook Han)
1980년 연세대학교 공과대학(공학사)
1987년 뉴욕공과대학교 대학원 컴퓨터공학과(공학석사)
1997년 오사카시립대학교 대학원 정보공학부(공학박사)
1998년~현재 가천대학교 소프트웨어학과 교수
관심분야 : 정보보호, 데이터 마이닝, 기계 학습
E-mail : mmhan@gachon.ac.kr