접근이 어려운 IOT 환경에서의 IDS를 위한 효과적인 특징 추출과 분류

이주화*, 박기현*
계명대학교 컴퓨터공학과
e-mail : yezi1004@gmail.com

Effective Feature Extraction and Classification for IDS in Accessible IOT Environment

Joo-Hwa Lee*, Ki-Hyun Park*
*Dept of Computer Engineering, Keimyung University

요 약

IOT는 복잡하고 이질적인 네트워크 환경이며 저전력 장치를 위한 새로운 라우팅 프로토콜의 존재로 인해 혁신적인 침입탐지 시스템이 필요하다. 특히 접근이 어려운 IOT 환경에서는 공격을 받았을 때 정확하고 빠른 탐지가 용이하여야 한다. 따라서 본 논문에서는 탐지의 정확성과 희소의 공격을 잘 탐지하기 위한 효과적인 특징 추출과 분류를 위한 SAR(Stacked Auto Encoder+Random Forest) 시스템을 제안한다.

1. 서론

현재의 네트워크는 예전에 비해 규모가 커지고 복잡성이 증가하고 있으며 특히, IOT는 엄청난 속도로 보급되어 2030년에는 약 125억 개에 이를 것으로 예상된다[1]. 또한, 비행 물체 등을 이용하여 센서들을 넓은 지역에 투하해서 배치시키고, 사람이 직접 유지보수하기 어려운 환경에서의 IOT가 증가하고 있다.

IOT는 무선, 유선, 위성, 셀룰러, 블루투스 등 다양한 유형에 따른 복잡성과 IOT 디바이스의 엄청난 수로 인하여 심각한 사이버 보안 위협과 취약성을 내포하고 있다[2]. 따라서 IOT 환경에서 침입탐지시스템(Intrusion Detection System)은 반드시 필요하다.

IDS는 탐지 방식에 따라서 오용탐지(Misuse Detection)방식과 비정상행위 탐지(Anomaly Detection) 방식으로 나눌 수 있다. 오용탐지는 미리 침입 행위를 정의하여 공격자의 공격 패턴이 탐지 시스템에 정의된 행위와 일치하면 침입 행위로 간주한다. 비정상행위 탐지는 어떤 행위가 정상적인 행동 패턴에서 어느 정도 벗어나면 침입 행위로 간주한다[3].

현재 IDS 연구가 많이 진행되고 있으나 침입이 다양해지고 새로운 형태가 나타나고 있기 때문에 단순한 특징 추출로 최신의 침입을 탐지할 수 없다. 또한 IDS의 성능은 특징의 선택에 따라 크게 달라지므로, 네트워크 트래픽을 정확하게 분류할 수 있는 특징을 선택하는 것은 여전히 어려운 과제이다[4]. 특히 IOT 데이터는 레이블이 없기 때문에 비지도 학습 모델을 사용하여야 한다.

따라서 본 연구에서는 IOT의 특성상 빠르고 정확하게 공격을 탐지를 할 수 있고 희소의 공격 탐지에도 효과적인 시스템을 제안하고자 한다.

2. 관련 연구

기존 연구에서는 오용탐지 기반의 네트워크 침입탐지 연구가 활발히 이루어졌으나 새로운 형태의 침입으로 인하여 많은 한계에 부딪히고 있다. 따라서 그 한계를 극복하기 위하여 비정상 탐지 기반으로 한 연구가 많이 진행되고 있다.

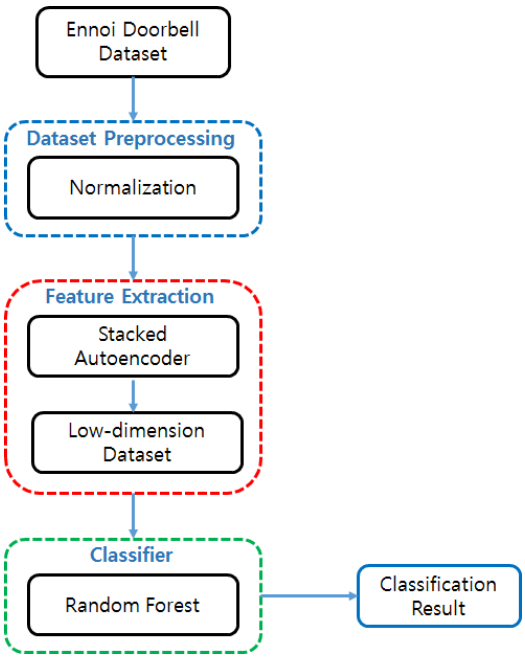
Prachi Shukla[5]는 IOT의 침입탐지를 위한 세 가지 머신러닝 알고리즘(K-means Clustering, Decision tree, hybrid two stage)을 사용하여 성능을 측정 비교하였다. 또한, Saeid Soheily-Khah[6] 등은 네트워크 침입 탐지를 위하여 지도학습과 비지도 학습을 결합한 하이브리드 침입 탐지 방법을 제안하였다. 하이브리드 침입탐지는 k-means clustering과 Random Forest를 결합하였으며 ISCX2012 데이터셋을 이용하여 성능을 측정하였다. Xiaoming Ye[7] 등은 네트워크 트래픽 특징 추출 및 분석하는 것이 IDS의 설계 및 구현의 기본으로 보고 저강도, 비정상 트래픽과 같이 트래픽 양에 변화를 나타내지 않는 비정상적인 트래픽을 탐지하기 위하여 Spark Streaming에 대한 접근 방식을 구현하였다. 트래픽 특징 추출 방법이 트래픽 변동과 통신 구조 변화를 효과적으로

※ 이 논문은 2018년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. NRF-2018R1D1A1B07043982).

탐지한다는 결과를 보여 주었다.

3. SAR 기반 IDS

본 논문에서 제안하는 접근이 어려운 IOT 환경에서의 특징 추출 및 분류 시스템은 그림 1과 같으며, 크게 데이터셋 전처리와 특징 추출, 분류 부분으로 나누어진다. 데이터셋의 전처리는 수치화와 정규화 순으로 처리하였다. 특징 추출은 Autoencoder 중 Stacked Autoencoder를 사용했으며 분류 알고리즘은 다수의 결정 트리들을 학습하고 다중 클래스 알고리즘 특성을 가지고 있는 Random Forest를 사용하였다.



(그림 1) SAR 기반 IDS

3.1 데이터셋

데이터셋은 UCI Machine Learning Repository에서 제공하는 IOT 데이터셋 중 Ennio Doorbell 데이터셋을 사용하였다. 정상 데이터와 분산된 서비스 거부 공격 (DDoS)을 시작하기 위해 Linux 시스템을 감염시키는 멀웨어 공격이 포함되어 있다.

훈련 데이터셋과 테스트 데이터셋은 <표 1>과 같이 각각 전체 데이터셋의 60%, 40%로 분할하여 사용하였다.

3.2 데이터셋의 전처리

데이터셋의 전처리의 정규화는 비교를 용이하게 하기 위하여 원래의 특징 값의 최대값과 최소값을 나타내는 최대 최소 정규화 방법을 사용한다.

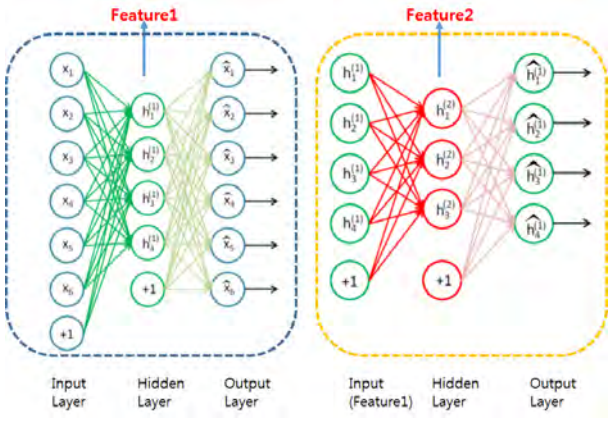
3.3 특징 추출

Autoencoder는 특정한 특징 벡터를 추상적인 특징 벡터로 점진적으로 변환할 수 있어 고차원 데이터 공간에서 저차원 데이터 공간으로의 비선형 변환을 할 수 있다. Autoencoder의 작업 프로세스는 인코딩과 디코딩의 두 단계로 나눌 수 있다. Autoencoder의 종류 중 Stacked Autoencoder는 여러 개의 히든 레이어를 가지는 Autoencoder이며, 레이어를 추가할수록 Autoencoder가 더 복잡한 부호화를 학습할 수 있다. Stacked Autoencoder는 (그림2)과 같이 가운데 히든레이어를 기준으로 대칭인 구조를 가진다.

히든레이어에서 특징을 추출하며 히든레이어의 수를 추가할수록 더 저차원의 특징을 추출할 수 있다. 차원이 줄어들수록 공격 탐지의 시간은 빠르게 된다.

<표 1> Ennio Doorbell 정상 및 공격 트래픽에 대한 훈련데이터와 테스트 데이터 수

Traffic		Training Data	Testing Data
Normal		23,460	15,640
Attack	combo	31,808	21,206
	junk	17,878	11,919
	scan	16,872	11,248
	tcp	60,921	40,615
	udp	62,359	41,573



(그림 2) Stacked Autoencoder를 이용한 특징 추출

3.4 분류

추출한 저차원의 특징을 기반으로 한 공격 탐지 시스템의 분류 알고리즘은 머신러닝의 대표적인 알고리즘인 Random Forest를 사용하였다. 머신러닝에서 여러 개의 모델을 학습시켜 그 모델들의 예측결과들을 이용해 하나의 모델보다 더 나은 값을 예측하는 방법을 앙상블 학습이라 하며 대표적인 예가 Random Forest이다. Random Forest는 여러 개의 Decision Tree들을 생성한 다음, 각각의 트리에서의 예측한 값 중에서 가장 많은 선택을 받은 클래스를 예측하는 알고리즘이다.

Decision tree는 학습데이터에 따라 생성되는 트리가 매우 달라지기 때문에 일반화가 쉽지 않다. 또한 계층적 접근 방식이므로 중간에 에러가 발생하면 다음 단계로 계속 에러를 전파하는 특성 때문에 IOT 환경에서는 적합한

분류 방법이 아니다.

이를 해결하기 위하여 Decision tree를 학습시킬 때 임의화 기술(Random space method)과 배깅을 사용하는 Random Forest 알고리즘으로 분류하였다. Random Forest는 임의성에 의해 서로 조금씩 다른 특성을 갖는 트리들로 구성되기 때문에 각 트리들의 예측들이 비 상관화되게 하며 결과적으로 일반화 성능을 향상 시킨다. 따라서 새로운 유형의 공격이나 희소의 공격 탐지에 효과적이다.

4. 실험 및 분석

실험 방법은 IOT 환경에서의 데이터 Ennoi Doorbell 데이터셋[8]과 기존 네트워크 환경의 데이터인 KDD99을 사용하여 본 논문에서 제안한 Stacked Autoencoder + Random Forest(SAR)로 분류한 결과를 비교하였다.

4.1 성능 평가

본 논문에서 결과를 측정하기 위하여 Confusion matrix에 기반한 메트릭을 사용한다. Confusion matrix의 정의는 <표2>에 나와 있다.

<표 2> Confusion matrix

		Actual	
		Positive	Negative
Predicted	Positive	TP	FP
	Negative	FN	TN

실험 결과는 수학식 1, 2, 3, 4 와 같이 Accuracy(정확도), Precision(정밀도), Recall(재현율), F1-score를 측정하였다.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

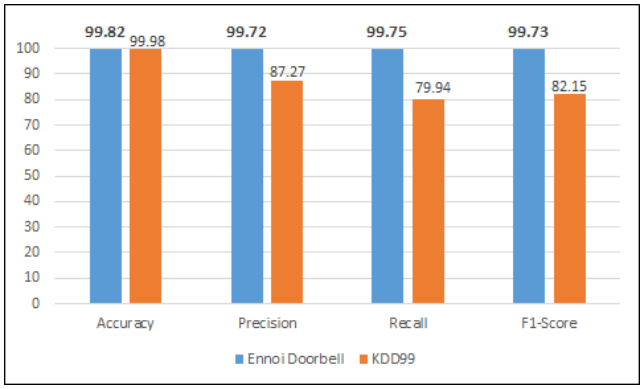
4.2 실험 결과

본 논문에서 제안한 SAR의 실험 결과는 <표3>과 같다. Ennoi Doorbell 데이터셋을 사용하여 SAR 모델로 실험하였을 때 전반적으로 성능이 98%이상으로 침입탐지에 효과적인 것으로 나타났다. 정확도뿐만 아니라 정밀도, 재현율, F1-Score의 성능이 모두 높게 나왔다.

<표 3> SAR 실험 결과

Class	Accuracy	Precision	Recall	F1-score
Normal	99.82%	99.95%	99.96%	99.96%
combo	99.82%	99.59%	99.37%	99.48%
junk	99.82%	98.89%	99.26%	99.07%
scan	99.82%	99.95%	99.98%	99.95%
tcp	99.82%	99.98%	99.98%	99.98%
udp	99.82%	99.98%	99.97%	99.97%

제안한 모델이 기존의 네트워크 환경에서도 적용이 되는지 확인하기 위하여 KDD99 데이터 셋과 비교하였다. 두 실험의 비교 결과는 (그림 3)과 같다.



(그림 3) IOT 환경의 데이터와 기존 네트워크 환경의 데이터 분류 성능 비교

Accuracy는 기존 네트워크 환경의 데이터에서 성능이 약간 높았으나 다른 성능은 모두 IOT 환경에서의 분류 성능이 우수하였다. IOT 환경에서의 데이터는 데이터 특성상 희소 클래스 문제가 많이 없으므로 특징 추출 및 분류에서 우수한 성능을 나타내었으나 기존 네트워크 환경에서의 데이터는 희소 클래스로 인하여 특징 추출과 학습이 제대로 되지 않은 R2L이나 U2R 클래스의 성능이 상대적으로 낮아 전체 성능이 낮게 나온 것을 볼 수 있다.

5. 결론 및 향후 연구

본 연구에서는 Stacked Autoencoder로 특징을 추출한 후 Random Forest로 분류하는 알고리즘을 제안하였다. Autoencoder는 딥러닝의 대표적인 비지도 학습 알고리즘으로 IOT의 특징을 추출하기 위한 적합한 모델이라 할 수 있다. 또한 수많은 네트워크의 고차원 특징을 저차원으로 압축한 후 분류를 하기 때문에 분류의 시간을 줄여준다. 따라서 본 연구에서 제안한 SAR 모델은 접근이 어려운 IOT 환경에서 효과적인 특징 추출과 분류에 적합한 모델이라 할 수 있다.

향후 연구 내용으로 IOT 환경의 공인 데이터셋 중 좀 더 많은 공격이 포함된 데이터셋에 적용할 것이며 기존

네트워크 환경에서의 데이터 중 IOT와 유사한 공격이 있는 데이터셋을 사용하여 새로운 유형의 공격 탐지를 실험하고자 한다. 또한 더욱 효과적인 침입탐지 시스템을 위하여 Autoencoder의 다른 모델과 여러 개의 파라미터를 조절하여 더욱 우수한 성능의 IDS를 제안하고자 한다.

참고문헌

[1] J. Howell, "Number of connected iot devices will surge to 125 billion by 2030 ihs markit says - ihs technology", [online] Available: https://technology.ihs.com/596542/.

[2] I. Alrashdi, A. Alqazzaz, E. Aloufi, R. Alharthi, M. Zohdy and H. Ming, "AD-IoT: Anomaly Detection of IoT Cyberattacks in Smart City Using Machine Learning", 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), pp. 305-310, 2019.

[3] B. YAN and G. HAN, "Effective Feature Extraction via Stacked Sparse Autoencoder to Improve Intrusion Detection System", IEEE ACCESS, vol6, pp. 41238-41248, 2018.

[4] F. Zhang and D. Wang, "An effective feature selection approach for network intrusion detection", 2013 IEEE Eighth International Conference on Networking, Architecture and Storage, Xi'an, pp. 307-311, 2013.

[5] P. Shukla, "ML-IDS: A machine learning approach to detect wormhole attacks in Internet of Things", 2017 Intelligent Systems Conference (IntelliSys), pp. 234-240, 2017.

[6] S. Soheily-Khah, P. Franc¸ois Marteau and N. B´echet, "Intrusion Detection in Network Systems Through Hybrid Supervised and Unsupervised Machine Learning Process: A Case Study on the ISCX Dataset", International Conference on Data Intelligence and Security, pp. 219-226, 2018.

[7] X. Ye, X.u Chen, D. Liu, W. Wang, L. Yang, G. Liang, and G. Shao, "Efficient feature extraction using apache spark for network behavior anomaly detection", Tsinghua Science and Technology, vol 23, 5, pp. 561-573, 2018.

[8] Machine Learning Repository. Accessed: Aug. 14, 2018. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/detection_of_IoT_botnet_attacks_N_BaIoT