JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 27 No. 7, pp. 101-108, July 2022
https://doi.org/10.9708/jksci.2022.27.07.101

Machine Learning-based Detection of DoS and DRDoS Attacks in IoT Networks

Seung-Yeon Yeo*, So-Young Jo*, Jiyeon Kim**

*Student, Dept. of Information Security, Seoul Women's University, Seoul, Korea
*Student, Dept. of Information Security, Seoul Women's University, Seoul, Korea
**Professor, Dept. of Computer Engineering, Daegu University, Gyeongsan, Korea

[Abstract]

We propose an intrusion detection model that detects denial-of-service(DoS) and distributed reflection denial-of-service(DRDoS) attacks, based on the empirical data of each internet of things(IoT) device by training system and network metrics that can be commonly collected from various IoT devices. First, we collect 37 system and network metrics from each IoT device considering IoT attack scenarios; further, we train them using six types of machine learning models to identify the most effective machine learning models as well as important metrics in detecting and distinguishing IoT attacks. Our experimental results show that the Random Forest model has the best performance with accuracy of over 96%, followed by the K-Nearest Neighbor model and Decision Tree model. Of the 37 metrics, we identified five types of CPU, memory, and network metrics that best imply the characteristics of the attacks in all the experimental scenarios. Furthermore, we found out that packets with higher transmission speeds than larger size packets represent the characteristics of DoS and DRDoS attacks more clearly in IoT networks.

▸ **Key words**: Internet of Things, Intrusion Detection, Machine Learning, Denial of Service, Distributed Reflection Denial of Service

[요 약]

본 논문은 다수의 사물인터넷 단말에서 보편적으로 수집할 수 있는 시스템 및 네트워크 메트릭을 학습하여 각 사물의 경험데이터를 기반으로 서비스거부 및 분산반사 서비스거부 공격을 탐지하는 침입 탐지 모델을 제안한다. 먼저, 공격 시나리오 유형별로 각 사물에서 37종의 시스템 및 네트워크 메트릭을 수집하고, 이를 6개 유형의 머신러닝 모델을 기반으로 학습하여 사물인터넷 공격 탐지 및 분류에 가장 효과적인 모델 및 메트릭을 분석한다. 본 논문의 실험을 통해, 랜덤 포레스트 모델이 96% 이상의 정확도로 가장 높은 공격 탐지 및 분류 성능을 보이는 것을 확인하였고, 그 다음으로는 K-최근접 이웃 모델과 결정트리 모델의 성능이 우수한 것을 확인하였다. 37종의 메트릭 중에는 모든 공격 시나리오에서 공격의 특징을 가장 잘 반영하는 CPU, 메모리, 네트워크 메트릭 5종을 발견하였으며 큰 사이즈의 패킷보다는 빠른 전송속도를 갖는 패킷이 사물인터넷 네트워크에서 서비스거부 및 분산반사 서비스거부 공격 특징을 더욱 명확히 나타내는 것을 실험을 통해 확인하였다.

▸ **주제어**: 사물인터넷, 침입탐지, 머신러닝, 서비스거부, 분산반사 서비스거부

• First Author: Seung-Yeon Yeo, Corresponding Author: Jiyeon Kim
*Seung-Yeon Yeo (downtime8454@swu.ac.kr), Dept. of Information Security, Seoul Women's University
*So-Young Jo (chan123@swu.ac.kr), Dept. of Information Security, Seoul Women's University
**Jiyeon Kim (jyk@daegu.ac.kr), Dept. of Computer Engineering, Daegu University
• Received: 2022. 06. 14, Revised: 2022. 07. 21, Accepted: 2022. 07. 21.

http://www.ksci.re.kr pISSN:1598-849X | eISSN:2383-9945

I. Introduction

사물인터넷(Internet of Things, 이하 IoT) 도입이 확산되면서 사이버 공격으로부터 IoT 환경을 보호하기 위한 IoT 보안 기술의 중요성이 높아지고 있다. IoT는 이기종의 물리적 개체들이 센서 데이터를 수집 및 전달하며 상호작용하는 네트워크이다[1]. 현재 네트워크에 연결된 단말 중, 가장 큰 비중을 차지하는 것이 IoT 단말이며 2025년경에는 네트워크에 연결된 IoT 단말이 약 300억 개를 초과할 것으로 예측된다[2]. 그러나 IoT 확산 속도에 비해 IoT 단말 및 네트워크에 적용되는 보안 기술 수준은 취약한 실정이다. 특히, IoT 트래픽의 98%가 암호화되어있지 않기 때문에 도청 및 남용의 위험성이 크고, IoT 단말의 약 57%가 심각한 위험성을 갖는 공격에 취약한 상태이다[3]. 또한, IoT 공격은 COVID-19로 인한 재택근무 증가 및 데이터 산업 활성화로 인해 2021년 상반기에는 전년 대비 100% 이상의 증가율을 보인 것으로 확인되었다[4]. 이와같이 IoT 공격이 급속히 확산되면서 IoT 환경에 최적화된 보안 기술을 개발하는 것이 필요하지만, IoT 네트워크에는 단말의 유형, 기능, 사양(specification), 플랫폼 등이 상이한 IoT 단말들이 혼재하기 때문에 전통적인 침입탐지시스템을 활용하여 규칙 또는 임계치 기반으로 IoT 공격을 탐지하는 것이 어렵다. 또한, 경량의 IoT 단말에는 공격 유형별로 여러 보안 솔루션을 설치 및 운영하는 것이 어렵다는 문제도 존재한다. 따라서 다수의 IoT 단말에서 보편적으로 수집할 수 있는 메트릭을 활용하여 각 단말의 경험데이터를 기반으로 공격을 탐지하는 침입탐지기술 개발이 필요하다.

최근 수행되는 침입탐지 연구에서는 규칙 및 임계치 기반의 전통적인 이상탐지 알고리즘을 개선하기 위하여 침입 데이터셋을 머신러닝(Machine Learning) 기반으로 학습하여 정상상태 및 공격상태를 모델링한다. 대표적인 침입 데이터셋으로는 침입탐지시스템의 성능평가를 위해 미국 방위 고등국에서 1999년 생성한 KDD CUP 1999[5]가 존재한다. 침입 데이터셋은 직접 다양한 공격을 주입하면서 방대한 트래픽을 수집 및 관리해야 하는 어려움으로 인해 공개된 데이터셋이 악성 소프트웨어(malware) 데이터셋에 비해 현저히 적으며 현재까지도 KDD CUP 1999 및 이를 개선한 NSL-KDD[6]이 침입탐지 연구를 위해 자주 사용되고 있다. 그러나 이들은 모두 전통적인 컴퓨터 네트워크에 공격을 주입하여 수집한 데이터셋이기 때문에 IoT 공격 탐지를 위해서 사용하는 데에는 한계가 있다. 대표적인 IoT 공격 데이터셋으로는 DS2OS[7], Bot-IoT[8], IoT-23[9] 등이 존재하며 이들은 전통적인 침입 데이터셋과 마찬가지로 네트워크 메트릭을 실시간 수집하여 데이터셋을 생성하였다. 또한, 서비스거부(Denial of Service, 이하 DoS) 공격, 분산 서비스거부(Distributed Denial of Service, 이하 DDoS) 공격 등의 네트워크 공격이 주요 IoT 공격으로 주입되었다. 이러한 공개 침입 데이터셋을 활용하여 IoT 보안 모델을 개발한 연구들이 다수 존재하지만, 이러한 보안 모델은 데이터셋에서 다루는 공격 및 공격 시나리오에 의존적이라는 한계가 존재한다. 또한, 공개 데이터셋에서 수집한 네트워크 메트릭 외에도 공격 탐지에 효과적인 메트릭을 발견하는 것도 침입탐지 연구에서 중요한 과제이다. 본 연구에서는 직접 IoT 네트워크를 구축하고 공격을 주입하면서 공격을 탐지하기 위해 필요한 메트릭 데이터를 실시간 수집한다. 단, 네트워크 메트릭뿐 아니라, 공격에 의한 증상을 관찰할 수 있는 IoT 단말의 시스템 메트릭도 함께 수집하고, 이들을 머신러닝 기반으로 학습하여 공격 탐지에 가장 효과적인 모델과 메트릭을 제안한다. IoT 공격으로는 DoS 공격, 그리고 공개 IoT 침입 데이터셋에 포함되지 않았던 분산반사 서비스거부(Distributed Reflection Denial of Service, 이하 DRDoS) 공격을 다양한 시나리오 하에 주입하여 데이터셋을 생성한다. DRDoS는 DDoS보다 발전된 공격으로서 반사체(reflector) 경유를 통해 공격자를 은닉하고 공격 증폭 효과를 누릴 수 있다. IoT 네트워크에서도 DRDoS가 발생될 수 있기 때문에 이를 효과적으로 탐지하기 위한 지능형 모델 개발이 필요하며 본 연구에서는 DRDoS 데이터셋을 직접 생성하여 머신러닝 기반 IoT 침입탐지 연구에 활용한다.

본 논문의 구성은 다음과 같다. 2장에서는 대표적인 IoT 공격 데이터셋 및 이를 활용한 관련 연구를 살펴보고, 3장에서는 본 연구의 데이터셋 수집 환경 및 공격 시나리오를 설명한다. 4장에서는 머신러닝 모델 유형 별 이진 및 다중분류 성능을 비교하고, DoS 및 DRDoS 탐지에 효과적인 메트릭을 분석한다. 5장에서는 결론 및 향후 연구를 제시한다.

II. Related Works

IoT 환경에서 수집된 대표적인 침입 데이터셋으로는 DS2OS[7], Bot-IoT[8], IoT-23[9], LITNET-2020[10], IoTID20[11], ToN_IoT[12] 등이 존재한다.

DS2OS(Distributed Smart Space Orchestration System)는 IoT 환경에서 침입탐지시스템의 이상탐지 알고리즘을 평가하기 위해 생성되었고, 조명 컨트롤러, 동작감지 센서, 온도 조절기, 태양열 배터리, 세탁기, 도어록

등의 IoT 단말로부터 IoT 트래픽의 송·수신지 주소, 접근 노드 주소 등을 데이터셋의 특징(feature)으로 수집하였다. Bot-IoT는 봇넷(botnet) 공격 데이터셋으로서 가상머신에서 IoT 봇넷 공격을 시뮬레이션하여 수집하였다. 데이터셋의 특징으로는 송·수신지 주소별 패킷 크기, 패킷 수 등을 수집한다[8]. IoT-23은 봇넷 소스코드를 이용하여 생성한 데이터셋으로서 라즈베리파이(RaspberryPi)에서 수집한 악성 트래픽과 스마트 램프, 인공지능 스피커 등에서 수집한 정상 IoT 트래픽으로 구성된다. IoT-23은 패킷의 프로토콜 유형, 패킷의 지속시간 등을 데이터셋의 특징으로 포함하며[9] LITNET-2020 데이터셋은 12개 유형의 공격을 주입하며 수집한 IoT 트래픽 데이터셋으로서 평균 패킷 크기 및 길이, 역방향 패킷의 길이 등을 데이터셋의 특징으로 포함한다[10]. IoTID20은 IoT 환경에서의 비정상 행위를 감지하기 위해 개발된 데이터셋으로서 노트북, 스마트폰, IP 카메라 등에서 패킷 지속시간, 초당 송·수신한 패킷 수, 평균 CPU 유휴 시간 등을 데이터셋의 특징으로 포함하여 데이터셋을 수집하였다[11]. ToN-IoT는 물리 시스템, 가상머신, 해킹 플랫폼, IoT 센서가 연결된 테스트베드 등을 활용하여 차세대 IoT 및 산업용 IoT 환경을 구축하고, 정상 및 악성 트래픽을 수집한 데이터셋이다. 데이터셋의 특징으로는 송·수신지의 총 패킷 수, 거부된 DNS(Domain Name Service) 쿼리, 디스크에서 읽어 들인 데이터의 양, 실행 중인 스레드(thread) 수 등을 포함한다[12].

위와 같은 공개 IoT 침입 데이터셋을 활용한 기존 IoT 보안 연구들은 주로 머신러닝 기반으로 데이터셋을 학습하고, 침입탐지에 효과적인 모델을 제안하였다. Sam Strecker[13] 및 Nicolas-Alin Stoian[14]은 IoT-23 데이터셋을 Random Forest(RF), Support Vector Machine(SVM), K-Nearest Neighbor(KNN), Naive Bayes(NB) 등의 머신러닝 모델을 활용하여 학습하고, RF 모델이 IoT 침입 탐지에 가장 효과적인 모델임을 실험을 통해 도출하였다. Nahida Islam[15]은 NSL-KDD, DS2OS, IoTID20 등의 데이터셋을 Decision Tree(DT), SVM과 같은 머신러닝 모델과 Deep Neural Network(DNN), Long Short-Term Memory(LSTM)과 같은 딥러닝(Deep Learning) 모델을 활용하여 학습하고, 머신러닝보다 딥러닝 모델이 공격 탐지에 더 효과적임을 보였다. Raneem Qaddoura[16]와 Hasan Alkahtani[17]는 IoTID20 데이터셋을 딥러닝 기반으로 학습하여 공격 탐지 및 공격 분류에 효과적인 모델을 제안하였고, Rawan Shahin[18]은 IoTID20 데이터셋의 일부 샘플에 대해 Logistic Regression(LR), Pearson's Correlation을 기반으로 중요한 특징을 추출하고, LR, RF, DT, KNN, NB, Adaboost 등의 머신러닝 모델을 기반으로 학습하여 Adaboost 및 RF 모델이 공격 탐지에 가장 효과적임을 보였다. 이 밖에도 DS2OS를 머신러닝 기반으로 학습한 연구[19], 머신러닝과 딥러닝 기반으로 학습한 결과를 비교하는 연구[20]가 존재하며 Bot-IoT 데이터셋을 머신러닝 및 딥러닝 기반으로 학습하는 연구[21-24]가 수행되었다.

기존에 공개된 IoT 공격 데이터셋은 DoS 공격 또는 DDoS 공격 샘플은 포함하지만, DRDoS 샘플은 포함하고 있지 않다는 한계가 있다. 또한, 네트워크 메트릭은 모든 데이터셋에서 수집하지만, 시스템 메트릭은 IoTID20 및 ToN-IoT 데이터셋만 수집하고 있다. 경량의 IoT 단말은 전통적인 컴퓨터 시스템에 비해 제한된 컴퓨팅 자원을 가지고 있기 때문에 공격 발생 시, 시스템 부하가 공격의 증상으로 나타날 수 있다.

본 연구에서는 네트워크 메트릭뿐 아니라, 시스템 메트릭을 DoS 및 DRDoS 공격을 주입하면서 실시간 수집하고, 이를 다양한 머신러닝 기반으로 학습하여 공격에 효과적인 모델을 제안한다. 2장에서 살펴본 대표적인 IoT 공격 데이터셋과 본 논문에서 수집하는 데이터셋의 메트릭 및 공격 유형을 비교하면 Table 1과 같다.

Table 1. Comparison of Existing IoT Intrusion Datasets and Our Dataset

Dataset	Metric			Attack		
	Net-works	Sys-tems	Sen-sors	DoS	DDoS	DR-DoS
DS2OS	O	-	O	O	-	-
Bot-IoT	O	-	-	O	O	-
IoT-23	O	-	-	O	-	-
IoTID20	O	O	-	O	-	-
LITNET-2020	O	-	-	O	-	-
ToN-IoT	O	O	-	O	O	-
Our Dataset	**O**	**O**	**-**	**O**	**-**	**O**

III. Collection of IoT Intrusion Datasets

본 장에서는 IoT 침입탐지 데이터셋을 직접 수집하고, 머신러닝 기반으로 분석하기 위한 실험 모델을 설계한다.

시스템 메트릭 및 네트워크 메트릭을 데이터셋의 특징으로 구성하는 IoT 침입 데이터셋을 생성하기 위하여 본 논문에서는 IoT 네트워크 공격 환경을 구축한 후, IoT 기

기에서 시스템 및 네트워크 메트릭을 정상상태 및 공격상태에서 실시간 수집한다. IoT 네트워크 공격으로는 IoT 환경에서 발생빈도가 높은 DoS 공격과 DDoS 공격이 진화된 DRDoS(Distributed Reflection DoS) 공격을 주입하여 생성한다. Fig. 1은 각 공격 주입을 위해 본 논문에서 구성한 실험환경을 보여준다.

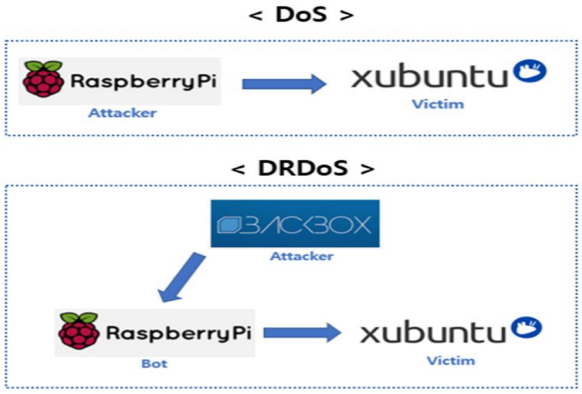

Fig. 1. System Deployments for DoS and DRDoS Injection

실험환경은 IoT 기기와 공격 및 피해 시스템으로 구성되며 IoT 기기로는 라즈베리파이를 사용하였다. DoS 공격에서는 IoT 기기가 대량의 공격 트래픽을 피해 시스템에 전송하고, DRDoS 공격에서는 공격 시스템이 공격 트래픽을 반사체인 IoT 기기를 통해 피해 시스템에 전송한다. 두 공격에서 모두 대량의 공격 트래픽이 피해 시스템에 전송되기 때문에 네트워크 및 시스템 자원이 고갈되어 결과적으로 DoS가 발생하게 된다.

DoS 및 DRDoS 공격 트래픽은 공격 강도에 따라 서로 다른 트래픽 패턴으로 생성될 수 있기 때문에 다양한 공격 시나리오를 고려하여 공격을 주입할 필요가 있다. 본 논문에서는 다양한 공격 시나리오를 고려하여 데이터셋을 생성할 수 있도록 Table 2와 같이 두 공격의 패킷 크기 및 속도를 조절하여 공격을 주입하며 데이터셋을 수집한다. 공격 데이터셋은 크게 정상(benign), 공격(attack) 트래픽으로 구분되고, 공격 트래픽은 패킷 크기와 패킷 전송속도가 각각 3단계로 설정된 시나리오로 생성된다.

Table 2. Experimental Scenarios of IoT DoS and DRDoS Attacks for Our Dataset Collection

Dataset	Scenario	bytes per packet	packets per second
D_{dos_size} D_{drdos_size}	benign	60	1
	attack1	100	1
	attack2	500	1
	attack3	1,000	1
D_{dos_speed} D_{drdos_speed}	benign	60	1
	attack4	60	10
	attack5	60	1,000
	attack6	60	6,000

정상 및 공격 데이터셋은 정상 샘플 1,200개와 공격 샘플 3,600개로 구성되며 각 샘플은 29종의 시스템 메트릭(메모리 11종, CPU 10종, 디스크 8종)과 8종의 네트워크 메트릭을 데이터셋의 특징으로 구성한다. Table 3은 37종의 시스템 및 네트워크 메트릭 중, 본 논문의 실험분석 시, 자주 등장하는 메트릭을 포함하여 대표적인 10개 메트릭에 대한 설명을 보여주고, Fig. 2는 37종의 특징으로 구성된 데이터셋의 예시로서 D_{dos_size} 및 D_{drdos_size}에 포함된 정상 및 공격 샘플을 보여준다.

Table 3. Key Features of Our IoT Intrusion Datasets consisting of System and Network Metrics

Metric		Description
Memory	kbbuffers	amount of memory buffers used by the kernel
	kbinact	amount of memory currently in use that has not been used recently
CPU	nice	CPU utilization at the user level
	soft	percentage of time spent processing software interrupts
Disk	rkB/s	size of data read from the disk per second
	wkB/s	size of data written to disk per second
Network	rxkB/s	packet size received per second
	txkB/s	packet size transmitted per second
	rxpck/s	number of packets sent per second
	txpck/s	number of packets transmitted per second

	1 usr(%)	2 nice(%)	3 sys(%)	4 iowait(%)	5 steal(%)	6 irq(%)	7 soft(%)	8 guest(%)	9 gnice(%)	10 idle(%)	11 tps	12 rkB/s	13 wkB/s	14 dkB/s	15 areq-SZ	16 aqu-SZ	17 await	18 util(%)	19 kbmemfree	20 kbavail
benign	50.52	39.18	10.31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10,4064	285,520
DoS	29.29	45.45	11.11	0	0	0	43.43	0	0	0	90	1,496	0	0	16.62	0.88	11.93	90.8	65,424	73,816
DRDoS	33	39	23	0	0	0	5	0	0	0	133	4,924	0	0	37.02	0	0.48	34.4	129,700	216,880

	21 kbmemused	22 memused(%)	23 kbbuffers	24 kbcached	25 kbcommit	26 commit(%)	27 kbactive	28 kbinact	29 kbdirty	30 rxpck/s	31 txpck/s	32 rxkB/s	33 txkB/s	34 rxcmp/s	35 txcmp/s	36 rxmcst/s	37 ifutil(%)
benign	446,504	45.76	43,172	235,620	3,703,808	153.62	328,496	273,232	356	1	1	0.06	0.06	0	0	0	0
DoS	661,568	67.8	9,740	95,868	3,713,792	154.04	315,296	317,836	132	6,400	6,335	375	371.19	0	0	0	0.31
DRDoS	511,140	52.39	17,496	167,684	3,498,100	145.09	368,624	194,464	372	5,918	0	346.76	0	0	0	0	0.28

Fig. 2. Example of Benign, DoS, and DRDoS Samples consisting of 37 Features

IV. Experimental Results

본 장에서는 3장에서 수집한 데이터셋을 6개의 머신러닝 모델(RF, KNN, LR, DT, SVM, NB)을 활용하여 학습하고, 공격 탐지 및 공격 강도 분류에 있어 F1-score가 높은 머신러닝 모델을 분석한다. F1-score는 모델의 정밀도(precision)와 재현율(recall)을 이용하여 계산한 조화평균의 값으로서 분류된 데이터의 수가 불균형을 이룰 때 사용하는 성능 지표이다. F1-score, 정밀도, 재현율은 수식 (1)과 같이 계산된다.

$$F1-score = 2\times \frac{(Recall * \mathrm{Precision})}{Recall + \mathrm{Precision}} \quad (1)$$

$$\text{단,} \quad \mathrm{Precision} = \frac{TP}{TP+FP} \text{ and } Recall = \frac{TP}{TP+FN}$$

TP(True Positive)는 정상 트래픽을 정상으로 올바르게 분류한 샘플 개수이고, FP(False Positive)는 공격 트래픽을 정상으로 잘못 분류한 샘플 개수, FN(False Negative)은 정상 트래픽을 공격으로 잘못 분류한 샘플의 개수를 의미한다. 본 논문에서는 각 데이터셋의 70%는 침입탐지 모델 개발을 위한 학습 데이터셋으로 사용하고, 30%는 개발된 모델의 성능을 측정하기 위한 평가 데이터셋으로 사용하였다.

본 장의 구성은 다음과 같다. 4.1장에서는 각 머신러닝 모델이 Table 4와 같이 공격 강도에 상관없이 정상 트래픽과 공격 트래픽으로 구분하는 능력, 즉, 네트워크에서 공격 트래픽을 정확하게 탐지해내는 이진분류(binary classification) 성능을 분석하고, 4.2장에서는 공격 탐지뿐 아니라, 각 머신러닝 모델이 공격 상황을 정확히 진단하는 능력, 즉, 공격 강도별로 공격 트래픽을 구분해내는 다중분류(multi-class classification) 성능을 분석한다.

Table 4. Types of Dataset Classification for Machine Learning of Our IoT Intrusion Datasets

Classification	Dataset	Label	Scenario
binary	D_{dos_size} D_{drdos_size}	benign	benign
		attack	attack1
			attack2
			attack3
	D_{dos_speed} D_{drdos_speed}	benign	benign
		attack	attack4
			attack5
			attack6
multi-class	D_{dos_size} D_{drdos_size}	benign	
		attack1	
		attack2	
		attack3	
	D_{dos_speed} D_{drdos_speed}	benign	
		attack4	
		attack5	
		attack6	

4.1 Binary Classification

D_{dos_size} 및 D_{drdos_size}를 머신러닝 모델을 활용하여 학습한 후, 공격 탐지 성능을 평가한 결과, Fig. 3(a)와 같이 RF 모델의 정확도가 가장 높고, Multinomial Naive Bayes(MNB) 모델의 정확도가 가장 낮은 것을 볼 수 있다.

MNB는 NB 모델의 한 유형으로서 본 논문에서는 MNB, Gaussian Naive Bayes(GNB), Bernoulli Naive Bayes(BNB) 모델을 사용하여 데이터셋을 학습하고, 이 중, 가장 정확도가 높은 모델을 NB 모델의 정확도로 반영한다. MNB 모델의 정확도가 낮은 이유는 두 공격 탐지에 있어 정상 트래픽을 공격 트래픽으로 잘못 분류하는 비율이 높았기 때문이다. 또한, KNN 모델을 제외한 5개 모델의 경우, DoS 탐지 정확도가 DRDoS 탐지 정확도에 비해 높은 것으로 나타났는데 이 역시 정상 트래픽을 공격 트래픽으로 잘못 탐지하는 비율이 DRDoS가 더 높았기 때문이다. 총 37종의 시스템 및 네트워크 메트릭 중, DoS 및 DRDoS 공격 트래픽을 탐지하는 데에 있어 가장 중요성이 높은 메트릭을 분석한 결과, 메모리 메트릭에 속하는 kbbuffers (커널에서 사용하는 메모리 버퍼의 양), 네트워크 메트릭에 속하는 rxkB/s(초당 수신 패킷 크기)가 공격으로 인한 증상을 가장 잘 반영하는 것으로 나타났다. 즉, DoS 및 DRDoS 공격 탐지를 위해서는 공격에 의해 증가하는 메모리 버퍼 크기 및 초당 수신되는 패킷 크기를 모니터링 하는 것이 효과적임을 실험을 통해 확인할 수 있다.

공격의 속도를 다르게 하여 수집한 D_{dos_speed} 및 D_{drdos_speed}를 머신러닝 모델을 활용하여 학습한 결과, Fig. 3(b)와 같이 RF, KNN, DT 모델의 탐지 정확도가 모두 0.95 이상으로 높고, MNB 모델은 가장 낮은 정확도를 보인 것을 확인할 수 있다. Fig. 3 그래프 (a)와 비교하면, (b)의 정확도가 대부분의 머신러닝 모델에서 높은 것을 볼 수 있는데 이는 DoS 및 DRDoS 공격의 경우, 패킷의 크기보다 패킷의 속도가 공격의 특징을 더욱 잘 반영한다고 해석할 수 있다. MNB 모델의 경우에는 그래프 (a)와 마찬가지로 6개의 모델 중, 정확도가 가장 낮았는데 (a)에서는 정상 트래픽을 공격 트래픽으로 잘못 분류하는 비율이 높았던 데에 반해 (b)에서는 공격 트래픽을 정상 트래픽으로 잘못 분류하는 비율이 높은 것으로 관찰되었다.

정상과 공격 트래픽을 분류하는 데에 있어서 가장 중요성이 높은 메트릭을 분석한 결과, 두 공격에서 모두, 메모리 메트릭 중에는 kbbuffers, kbinact(현재 사용하는 메모리 중, 최근에 사용되지 않았던 메모리의 양), CPU 메트릭 중에는 nice(사용자 레벨에서의 CPU 사용률), 네트워크 메트릭 중에는 txpck/s(초당 전송된 패킷 수)가 두 공격으로 인한 증상을 가장 잘 반영하는 것으로 나타났다.

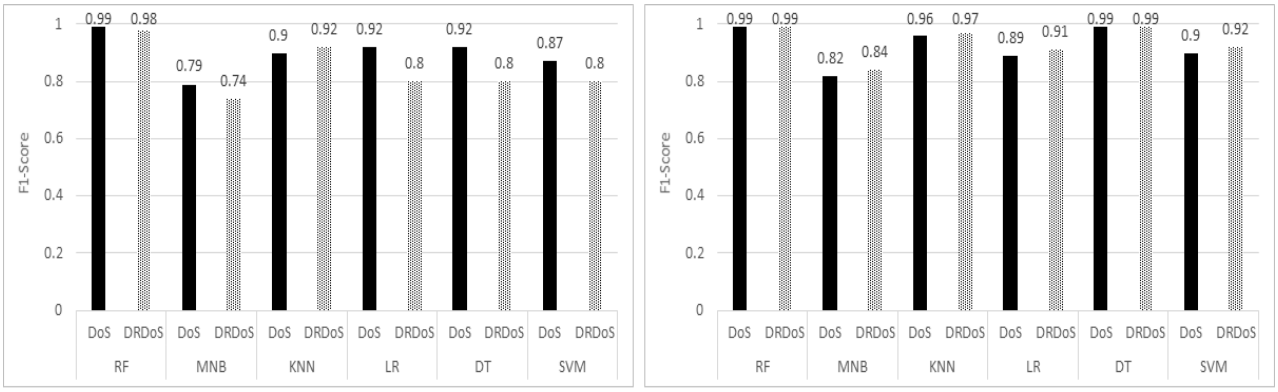

Fig. 3. Performance of Binary Classification by Machine Learning Models – (a) training D_{dos_size} and D_{drdos_size}, (b) training D_{dos_speed} and D_{drdos_speed}

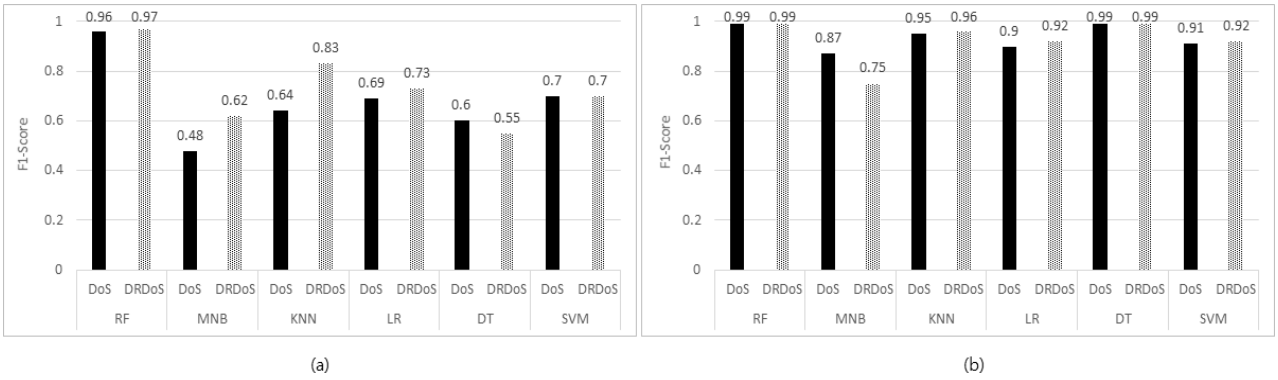

Fig. 4. Performance of Multi-class Classification by Machine Learning Models – (a) training D_{dos_size} and D_{drdos_size}, (b) training D_{dos_speed} and D_{drdos_speed}

4.2 Multi-class Classification

본 논문에서 수집한 데이터셋에 대해 머신러닝 기반의 다중분류 연구를 수행하면 정상 트래픽과 공격 트래픽을 구분하는 성능 외에도 공격 트래픽을 공격 강도에 따라 분류해내는 성능을 평가할 수 있다. Fig. 4(a)는 D_{dos_size} 및 D_{drdos_size}에 대한 6개의 머신러닝 모델 유형별 다중분류 성능을 보여준다. 이진분류 결과와 마찬가지로 RF 모델의 성능이 가장 높고, 나머지 5개 모델은 Fig.3(a)의 이진분류 성능(F1-score 0.74-0.92)에 비해 현저하게 낮은 정확도(F1-score 0.48-0.83)를 보이는 것을 확인할 수 있다. MNB 모델은 6개의 머신러닝 모델 중에는 가장 낮은 성능을 보이지만 GNB 및 BNB 모델보다는 높기 때문에 NB 모델의 대표 값으로 사용하였다. Fig.4(b)는 D_{dos_speed} 및 D_{drdos_speed}에 대한 다중분류 성능 결과를 보여준다. Fig. 3(b)의 이진분류 결과와 마찬가지로 RF와 DT의 성능이 0.99로 가장 높았고, 그 다음으로는 KNN, LR, SVM이 0.9 이상으로 좋은 성능을 보였다.

Fig. 4의 (a)와 (b)를 비교하면, 공격 패킷 크기가 다양한 공격들을 서로 구분하는 것이 패킷 전송속도가 다양한 공격들을 서로 구분해내는 것보다 어렵기 때문에 (a) 정확도가 (b) 정확도보다 더 낮은 것을 확인할 수 있다.

Fig. 4(a)의 DRDoS 다중분류 결과를 Fig. 5와 같이 혼동행렬(confusion matrix)을 통해 상세 분석한 결과, RF 모델을 제외한 5개 머신러닝 모델은 정상 시나리오, attack1 및 attack2 시나리오 간의 오분류율이 높아 정확도가 낮은 것을 알 수 있다. 반면, attack3 시나리오는 공격 강도가 가장 높기 때문에 모든 모델에서 분류 정확도가 가장 높은 것을 볼 수 있다.

마지막으로 다중분류에 있어 가장 중요성이 높은 메트릭을 분석한 결과, 모든 데이터셋에서 메모리 메트릭인 kbbuffers, kbinact가 공격 강도에 따른 증상을 가장 잘 반영하는 것으로 분석되었고, 추가적으로 DoS 공격에서는 rxkB/s, DRDoS 공격에서는 rxpck/s, nice 메트릭이 공격 증상을 잘 반영하는 것으로 확인되었다. 즉, 공격 발생 여부뿐 아니라, 공격 강도에 따른 정확한 트래픽 진단을 위해서는 위 메트릭들을 모니터링 하는 것이 효과적임을 실험을 통해 확인할 수 있었다.

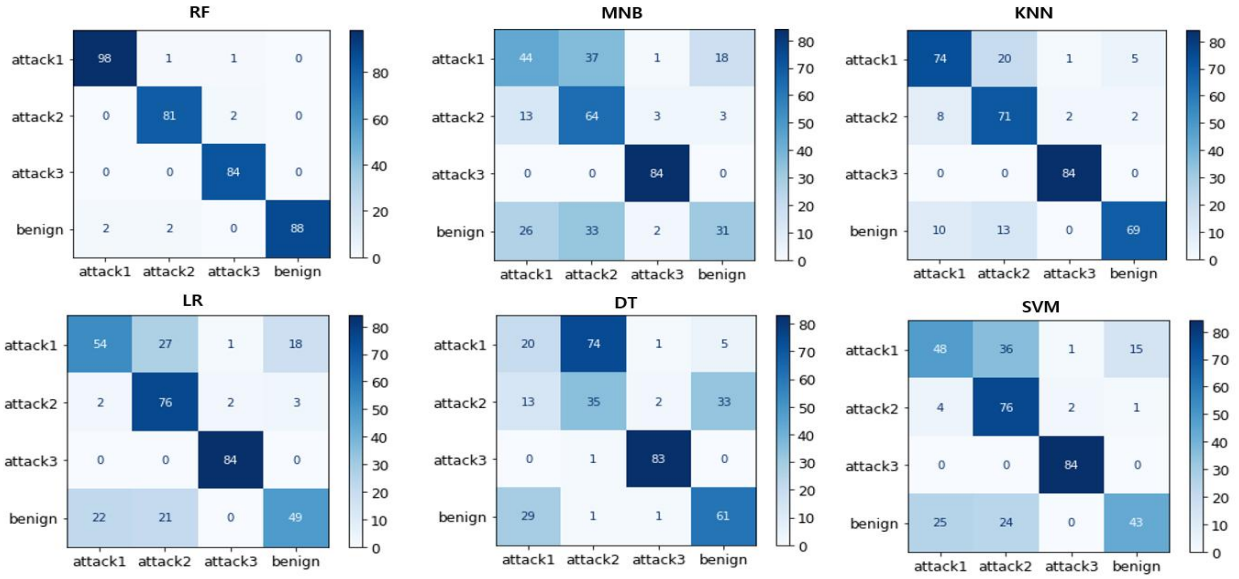

Fig. 5. Confusion Matrix of Multi-class Classifications for D_{drdos_size}

V. Conclusion

본 논문에서는 IoT 네트워크에서 발생하는 DoS 및 DRDoS 공격을 머신러닝 기반으로 탐지하는 침입탐지 모델을 개발하기 위하여 공격 강도(패킷 크기, 패킷 속도)를 다양하게 하여 공격을 주입하고, 총 37종의 시스템 및 네트워크 메트릭을 실시간 수집하여 IoT 침입 데이터셋을 생성하였다. 생성된 데이터셋을 RF, KNN, LR, DT, SVM, NB와 같은 6개 유형의 머신러닝 모델을 활용하여 학습한 결과, 모든 공격 시나리오에서 RF 모델이 이진분류(공격 탐지) 및 다중분류(정상 및 공격 강도 세부 분류)에 가장 좋은 성능을 보이는 것으로 분석되었다. 또한, 공격 패킷의 크기보다 공격 패킷을 전송하는 속도가 공격 특징을 더욱 잘 반영하기 때문에 머신러닝 기반의 공격 탐지율이 더 높은 것으로 나타났고, 모든 공격 시나리오에서 메모리 메트릭인 kbbuffers, kbinact 메트릭이 DoS 및 DRDoS 공격 증상을 가장 잘 반영하는 메트릭으로 확인되었다. 이 밖에도 네트워크 메트릭 중에는 rxkB/s가 DoS 공격 탐지에 효과적이고, rxpck/s는 DRDoS 탐지에 효과적인 것으로 나타났고, CPU 메트릭 중에서는 nice 메트릭이 공격 속도에 따른 DoS 및 DRDoS 공격 탐지에 효과적임을 확인하였다.

향후에는 본 논문에서 식별한 중요 메트릭들을 활용하여 경량이면서도 정확도가 높은 IoT 침입탐지 솔루션을 구현할 예정이며 6개 유형의 머신러닝 모델 외에도 다양한 딥러닝 모델을 활용하여 DoS 및 DRDoS 공격 탐지에 효과적인 모델을 개발하는 연구를 수행할 계획이다.

REFERENCES

[1] Patel Keyur, Patel Sunil Scholar P, and Salazar Carlos, "Internet of Things-IOT: Definition, Characteristics, Architecture, Enabling Technologies, Application & Future Challenges," IJESC, Vol. 6, NO. 5, May 2016. DOI: 10.4010/2016.1482

[2] Knud Lasse Lueth, "State of the IoT 2020: 12 billion IoT connections, surpassing non-IoT for the first time", https://iot-analytics.com/state-of-the-iot-2020-12-billion-iot-connections-surpassing-non-iot-for-the-first-time

[3] Paloaltonetworks,"2020 Unit 42 IoT Threat Report", https://unit42.paloaltonetworks.com/iot-threat-report-2020, 2020.03.10.

[4] IoTnews, “Kaspersky: Attacks on IoT devices double in a year”, https://www.iottechnews.com/news/2021/sep/07/kaspersky-attacks-on-iot-devices-double-in-a-year, 2021.09.07

[5] KDD Cup 1999 Data, http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[6] NSL-KDD dataset, https://www.unb.ca/cic/datasets/nsl.html

[7] FrancoisXA, DS2OS traffic traces, https://www.kaggle.com/francoisxa/ds2ostraffictraces

[8] Koroniotis Nickolaos, Moustafa Nour, Sitnikova Elena, and Turnbull Benjamin, “Towards the Development of Realistic Botnet Dataset in the Internet of Things for Network Forensic Analytics: Bot-IoT Dataset,” Future Generation Computer Systems, Volume 100, pp. 779-796, Nov. 2019.

[9] Sebastian Garcia, Agustin Parmisano, and Maria Jose Erquiaga, IoT-23: A labeled dataset with malicious and benign IoT network traffic, https://www.stratosphereips.org/blog/2020/1/22/aposemat-iot-23-a-labeled-dataset-with-malicious-and-benign-iot-network-traffic

[10] Damasevicius Robertas, Venčkauskas Algimantas, Grigaliunas Sarunas, Toldinas Jevgenijus, Morkevicius Nerijus, Aleliūnas Tautvydas, and Smuikys Paulius,“ LITNET-2020: An Annotated Real-World Network Flow Dataset for Network Intrusion Detection,” Electronics, Vol. 9, No. 5, 2020. DOI: 10.3390/electronics9050800

[11] Ullah Imtiaz, and Mahmoud Qusay. “A Scheme for Generating a Dataset for Anomalous Activity Detection in IoT Networks,” Advances in Artificial Intelligence, pp. 508-520, May 2020. DOI: 10.1007/978-3-030-47358-7_52

[12] Booij Tim, Chiscop Irina, Meeuwissen Erik, Moustafa Nour, and den Hartog Frank, “ToN_IoT: The Role of Heterogeneity and the Need for Standardization of Features and Attack Types in IoT Network Intrusion Datasets,” IEEE Internet of Things Journal, Vol. 9, NO. 1, pp. 485-496, Jan 2022. DOI: 10.1109/JIOT.2021.3085194

[13] Strecker Sam, Dave Rushit, Siddiqui Nyle, and Seliya Naeem, “A Modern Analysis of Aging Machine Learning Based IoT Cybersecurity Methods,” Journal of Computer Sciences and Applications, Vol. 9, NO. 1, pp. 16-22, Oct 2021. DOI: 10.12691/jcsa-9-1-2

[14] Nicolas-Alin Stoian, “Machine Learning for Anomaly Detection in IoT networks: Malware analysis on the IoT-23 Data set,” Jul 2020.

[15] Islam Nahida, Farhin Fahiba, Sultana Ishrat, Kaiser M. Shamim, Rahman Md, Hosen A. S. M., Cho Gi, and Hwan Gi, “Towards Machine Learning Based Intrusion Detection in IoT Networks,” Computers, Materials and Continua, Vol. 69, NO. 2, pp. 1801-1821, Aug 2021. DOI:10.32604/cmc.2021.018466

[16] Raneem Qaddoura, Ala’M. Al-Zoubi, Hossam Faris, and Iman Almomani, “A Multi-Layer Classification Approach for Intrusion Detection in IoT Networks Based on Deep Learning”, Sensors, Vol. 21, NO. 9, Apr 2021. DOI: 10.3390/s21092987

[17] Hasan Alkahtani, and Theyazn H. H. Aldhyani, "Intrusion Detection System to Advance Internet of Things Infrastructure-Based Deep Learning Algorithms", Complexity, Vol. 2021, NO. 3, Jul 2021. DOI: 10.1155/2021/5579851

[18] Shahin Rawan, and Sabri Khair Eddin, “A Secure IoT Framework Based on Blockchain and Machine Learning,” International Journal of Computing and Digital Systems, Vol. 11, NO. 1, pp. 671-683, Jan 2022. DOI: 10.12785/ijcds/110154

[19] Hasan Mahmudul, Islam Md, Islam Ishrak, and Hashem M.M.A., “Attack and Anomaly Detection in IoT Sensors in IoT Sites Using Machine Learning Approaches,” Internet of Things, Sep 2019. DOI: 10.1016/j.iot.2019.100059

[20] Reddy Dukka, Behera Dr. H., Nayak Janmenjoy, Vijayakumar P, Naik Bighnaraj, and Singh Pradeep, “Deep neural network based anomaly detection in Internet of Things network traffic tracking for the applications of future smart cities,” Transactions on Emerging Telecommunications Technologies, Vol. 32, NO. 6, Oct 2020. DOI: 10.1002/ett.4121

[21] Shafiq Muhammad, Tian Zhihong, Sun Yanbin, Du Xiaojiang, and Guizani Mohsen, “Selection of effective machine learning algorithm and Bot-IoT attacks traffic identification for internet of things in smart city,” Future Generation Computer Systems, Vol. 107, NO. 4, Jun 2020. DOI: 10.1016/j.future.2020.02.017

[22] Satish Pokhrel, Robert Abbas, and Bhulok Aryal “IoT Security: Botnet detection in IoT using Machine learning,” arXiv:2104.02231, Arp 2021. DOI: 10.48550/arXiv.2104.02231

[23] Churcher Andrew, Ullah Rehmat, Ahmad Jawad, Rehman Sadaqat Ur, Masood Fawad, Gogate Mandar, Alqahtani Fehaid, Nour Boubakr, and Buchanan William, “An Experimental Analysis of Attack Classification Using Machine Learning in IoT Networks,” Sensors, Vol. 21, NO. 2, pp. 1-32, Jan 2021. DOI: 10.3390/s21020446

[24] Das Anurag, Ajila Samuel, and Lung Chung-Horng, “A Comprehensive Analysis of Accuracies of Machine Learning Algorithms for Network Intrusion Detection,” Machine Learning for Networking, pp. 40-57, Apr 2020. DOI: 10.1007/978-3-030-45778-5_4

Authors

Seung-Yeon Yeo is an undergraduate student in the Department of Information Security at Seoul Women’s University, Seoul, Korea, since 2019. Her research interests include cybersecurity, internet of things, and artificial intelligence.

So-Young Jo is an undergraduate student in the Department of Information Security at Seoul Women’s University, Seoul, Korea, since 2017. Her research interests include cybersecurity, internet of things, and artificial intelligence.

Jiyeon Kim received the B.S. and Ph.D. degrees in information security engineering from Seoul Women’s University, Seoul, Korea, in 2007 and 2013, respectively. She was a Postdoctoral Research Associate in the Department of Electrical and Computer Engineering, Carnegie Mellon University, United States, from 2014 to 2017. She is currently a Professor in the Department of Computer Engineering, Daegu University, Gyeongsan, Korea. Her research interests include cybersecurity, cloud computing, internet of things, artificial intelligence, and critical infrastructure protection.