상용망에 적합한 머신러닝 기반 DDoS 탐지 시스템

임호문
고려대학교 소프트웨어보안학과 석사과정
limha25@naver.com

Machine Learning-Based DDoS Detection System Suitable for Commercial Networks

Ho-Mun Lim
Department of Software Security, Korea University

요 약

분산 서비스 거부(DDoS) 공격은 대규모 네트워크 운용 환경에서 탐지 및 대응이 지연될 경우 심각한 피해를 초래할 수 있는 중대한 보안 위협이다. 이러한 위협을 효과적으로 방어하기 위해 규칙기반 탐지, 머신러닝, 딥러닝 등을 활용한 다양한 탐지 기법들이 활발히 연구되고 있다. 본 연구에서는 기존 공개된 실험망 데이터 기반의 DDoS 탐지 연구 한계를 극복하기 위해 상용망에서 수집된 실제 데이터를 활용하여 연구하였다. 또한, 다양한 머신러닝 알고리즘을 활용하여 DDoS 탐지 성능을 비교·분석하고, 이를 통해 상용망 환경에 최적화된 머신러닝 기반 DDoS 탐지 시스템을 제안한다

1. 서론

분산 서비스 거부(Distributed Denial of Service, DDoS) 공격은 대량의 악성 트래픽을 목표 시스템에 전송함으로써 해당 시스템의 서비스를 중단시키거나 가용성을 저하시키는 심각한 보안 위협이다. 대규모 네트워크에서 DDoS 공격의 탐지와 대응이 지연되면 심각한 서비스 중단이 발생할 수 있기 때문에, 신속하고 정확한 탐지 및 대응이 매우 중요하다.

기존의 지식기반 및 규칙기반 DDoS 탐지 알고리즘은 해커의 날로 지능화되는 공격을 탐지하지 못하는 한계가 있으며, 차단이 지연되는 문제가 발생할 수 있다. 또한, 인터넷 사용 환경의 고사양화, 대용량화 및 고속화로 발생하는 과다 트래픽을 공격트래픽으로 잘못 판단해 정상 트래픽을 차단하는 오류를 일으킬 수 있다[1].

이러한 다양한 문제를 해결하기 위해 보안 전문가의 실시간 탐지와 대응이 필요하나 대규모 네트워크 운용 환경에서 모든 DDoS 공격을 보안 전문가가 대응하는 것은 시간, 인력 부족으로 어려운 실정이다[1].

최근 여러 연구자들이 머신러닝, 딥러닝 기술을 활용하여 DDoS 공격을 탐지하고 차단하는 다양한 메커니즘을 연구 및 제안하고 있다[2]. 그러나 이러한 탐지 방법론들은 여전히 한계와 문제점을 가지고 있다. 예를 들어, 불안정한 트래픽 패턴, 익명화된 데이터, 이상치·결측치 데이터, 오래된 공격시나리오 등으로 인해 탐지, 방어 모델을 테스트하고 평가할 적절한 데이터 세트를 확보하는 데 어려움이 있다[2].

또한 대부분의 연구는 공개된 실험망 환경에서 생성된 데이터에 의존하고 있어 실제 상용망에서 검증이 부족한 실정이다.

따라서 본 연구에서는 캐나다 University of New Brunswick(UNB)에서 제공하는 CIC-DDoS2019 데이터 세트와 상용망에서 수집한 실제 데이터를 활용하여 상용망에 적합한 머신러닝 기반의 실시간 DDoS 탐지 시스템을 제안한다.

2. 관련연구

2.1 분산 서비스 거부 공격 탐지 방법

DDoS 공격을 탐지하기 위한 다양한 방법론이 제안되어 왔다. 초기의 DDoS 탐지시스템은 사전에 정의된 공격 패턴과 특정 규칙에 따라 DDoS 트래픽을 탐지하는 방식을 사용하였다[3]. 이러한 규칙 기반 접근 방식은 특정 패턴에 의존하기 때문에 최근의 지능화된 새로운 형태의 공격을 탐지하는 데 한계가 존재한다. 최근에는 Random Forest, CatBoost, XGBoost 등 머신러닝과 딥러닝을 활용한 탐지방법론이 활발히 연구되고 있다[4]. 머신러닝 기반의 DDoS 탐지 시스템은 정상 트래픽과 비정상 트래픽의 패턴을 학습하여 새로운 공격형태에 적응할 수 있는 강점이 있다.

2.2 DDoS 탐지 연구 데이터 세트

머신러닝 기반 DDoS 탐지 연구에 사용되는 대표적 데이터 세트로는 NSL-KDD, CAIDA, UNSW-NB15, Bot-IoT, CICIDS2017, CIC-DDoS2019 등이 있다[5]. NSL-KDD 는 정상트래픽과 다양한 네트워크 공격이 포함된

데이터 세트를 제공하며, CAIDA 의 "DDoS 공격 2007" 데이터 세트는 DDoS 공격과 관련된 익명화된 트래픽을 포함하고 있다. UNSW-NB15 는 Fuzzers, Analysis, Exploits 등 9 가지 공격 유형과 49 개의 피처를 제공하며, Bot-IoT 는 IoT 장치에서 발생 하는 정상 트래픽과 봇넷의 공격 트래픽을 포함하고 있다. CICIDS 2017 데이터 세트는 DDoS 공격, 웹공격, 침입공격 등으로 구성되어 있으며 CIC-DDoS 2019 는 가장 최근의 DDoS 공격 샘플을 포함하고 있다.

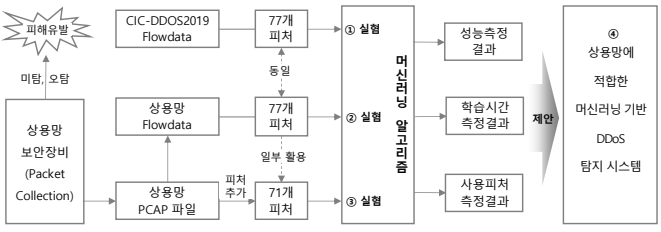
2.3 머신러닝 알고리즘

DDoS 탐지에 주로 사용되는 머신러닝 알고리즘에는 Random Forest, CatBoost, Light GBM, XGBoost, GBM 등이 있다[4].

Sharafaldin[2]의 연구에서는 CIC-DDoS2019 데이터 세트에서 Flow 기반으로 피처(feature)를 추출하여 ID3, Random Forest, 로지스틱 회귀(Logistic Regression), 나이브 베이즈(NaiveBayes) 알고리즘을 사용해 성능 평가를 수행하였다. ID3 알고리즘은 의사결정트리를 구성하기 위해 데이터 집합을 재귀적으로 분할하고, 최적의 속성을 찾기 위해 엔트로피 개념을 사용한다. Random Forest는 의사결정트리와 앙상블 학습 기법을 결합한 머신러닝 알고리즘이다. 세 가지 평가 지표(정밀도, 재현율, F1 Score)의 가중 평균 결과에 따르면, Random Forest와 ID3 알고리즘 이 높은 정확도를 달성하였으며, 재현율 측면에서는 ID3 알고리즘이 우수한 성능을 보였다.

3. 방법론

본 연구에서는 보안장비에서 발생하는 DDoS 공격 미탐지(False Negative)와 오탐지(False Positive)를 개선하기 위해 그림 1과 같이 실험 환경을 구성하고 다양한 머신러닝 알고리즘을 연구, 실증하여 최적의 DDoS 탐지 시스템을 제안한다.



(그림 1) DDoS 탐지 연구 방법

본 연구에서 DDoS 탐지 연구방법은 다음과 같다

첫 번째, UNB에서 제공하는 CIC-DDoS2019 데이터 세트의 Flow 기반 피처를 활용하여 다양한 머신러닝 알고리즘의 성능과 학습시간을 분석한다.

두 번째, 상용망에서 수집한 PCAP 파일에서 첫 번째 CIC-DDoS2019 데이터 세트와 동일한 Flow 기반 피처를 추출하여 머신러닝 알고리즘의 성능과 학습시간을 분석한다.

셋째, 상용망에서 수집한 PCAP 파일기반 분석을 위해 두번째 단계의 피처에서 불필요한 피처를 제거하고, 파일 기반의 일부 피처를 추가하여 머신러닝 알고리즘 성능과 학습시간을 분석 한다.

마지막으로, 상용망 환경에 적합한 머신러닝 기반의 실시간 DDoS 탐지 시스템을 제안한다.

4. 실험환경

본 연구의 머신러닝 실험은 Python 3.12 환경에서 수행되었으며, 데이터처리 및 모델구현을 위해 Numpy, Pandas, Scikit-learn, XGBoost, LightGBM, CatBoost, Matplotlib, RandomizedSearchCV 등 다양한 라이브러리를 사용하였다.

실험 데이터는 CIC-DDoS2019 Kaggle 데이트 세트와 상용망에서 수집한 PCAP 데이터 세트를 활용한다.

4.1 데이터 세트

4.1.1 CIC-DDoS2019 데이터 세트

CIC-DDoS2019 데이터 세트는 캐나다 University of New Brunswick(UNB)에서 제공하는 Flow 기반 데이터 세트로 다양한 최신 DDoS 공격 유형을 포함하고 있다[2]. 이 데이터 세트는 이상치, 결측치가 포함되어 전처리가 필요하다. 이를 위해 본 연구에서는 Kaggle 에서 제공하는 클리닝된 Parquet 파일을 사용하였으며, 해당 데이터는 11 개 Parquet 파일과 77 개 피처로 구성되어 있다.

4.1.2 상용망 데이터 세트

상용망 PCAP 데이터 세트는 2024 년 7 월 1 일부터 8 월 15 일까지 상용망에서 수집한 DDoS 트래픽 PCAP 파일 824 개와 정상 트래픽 PCAP 파일 788 개로 구성하였다. 이를 바탕으로 두 가지 유형의 데이터 세트를 마련하였다. 첫 번째 유형은 CIC-DDoS2019 데이터 세트와 동일한 방식으로 77 개의 Flow 기반 피처를 추출한 데이터 세트이다. 두번째 유형은 PCAP 파일 기반의 데이터인데, Flow 기반 데이터는 데이터 수가 많아 시스템 성능 저하가 발생될 수 있으므로 파일 기반 데이터 세트를 구성하였다. CIC-DDoS2019 데이터 세트에서 제공하는 일부 피처를 제거하고 PCAP 파일에서만 추출할 수 있는 출발지 IP, 목적지 IP 엔트로피 등의 추가적인 피처를 포함하여 71 개 피처로 구성하였다[6].

4.2 데이터 전처리

모델 학습을 위해서는 데이터의 품질을 향상시키기 위한 전처리가 필수적이다. 모든 데이터세트는 결측값(NaN)은 '0'으로 처리하였고, 문자형 데이터는 숫자형(float)으로 변환하였다.

또한, 상용망 데이터 세트는 피처정규화, 스케일링

등을 통해 머신러닝 알고리즘 학습 효율성을 높였고, StandardScaler 를 사용하여 피처 스케일링을 진행하였다. 데이터 불균형 문제를 해소하기 위해서 ADASYN(Adaptive Synthetic Sampling) 기법을 사용하여 소수 클래스의 데이터를 증강하였다.

4.3 k-fold 교차검증 및 하이퍼파라미터 최적화

모델 성능 평가를 위해 k-fold 교차검증(K=10)을 사용하였다. 이 방법은 데이터 세트를 10개 폴드로 나누어 각 폴드에서 교차 검증을 통해 모델의 성능을 평가한다. 또한, 각 모델의 최적화된 성능을 도출 하기 위해 RandomizedSearchCV를 사용하여 하이퍼 파라미터 튜닝을 진행하였다. RandomizedSearchCV는 랜덤하게 샘플링하여 최적 파라미터를 탐색하는 방식으로 효율적인 최적화를 가능하게 한다.

모델 성능평가는 Accuracy, Precision, Recall, F1 Score를 사용하여 평가하였다

5. 연구결과

5.1 CICDDoS2019 Flow 기반 데이터세트 연구결과

CIC-DDoS2019 Flow 기반 데이터세트를 활용한 성능평가 결과, 모든 알고리즘이 99.9% 이상의 정확도를 기록하였으며, 그 중에서도 Random Forest가 가장 높은 정확도를 보였다. 학습 시간 측면에서는 Light GBM과 XGBoost 알고리즘이 가장 빠른 속도를 기록하였다. 성능과 학습시간을 종합적으로 고려시 Light GBM과 XGBoost가 DDoS 탐지에 가장 적합한 알고리즘으로 평가된다.

표 1은 CIC-DDoS2019 데이터 세트에 대한 각 알고리즘의 성능평가 결과를 나타낸다.

구분	Accuracy	Precision	Recall	F1 Score	Time(s)
LightGBM	0.9998	0.9996	0.9995	0.9995	17.8
XGBoost	0.9998	0.9997	0.9996	0.9996	23.7
랜덤포레스트	0.9999	0.9998	0.9999	0.9998	331.9
CatBoost	0.9997	0.9992	0.9993	0.9992	370.1
GBM	0.9987	0.9962	0.9981	0.9971	1202.1

<표1> CIC-DDoS2019 Flow 기반 데이터 세트 평가 결과

5.2 상용망 데이터 세트 연구결과

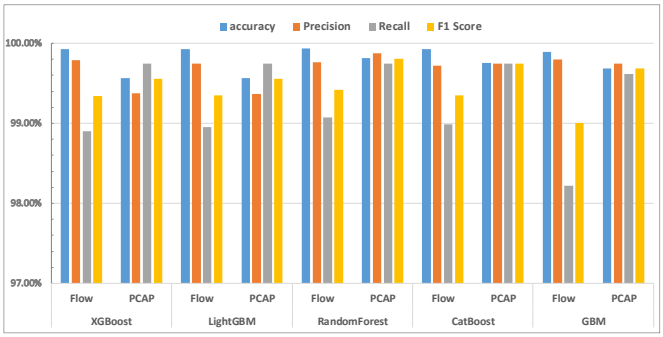
2가지 유형의 데이터 세트를 활용하여 연구하였다.

첫번째로, 상용망 Flow 기반 데이터 세트의 성능평가 결과, 모든 알고리즘이 99% 이상의 높은 정확도를 기록하였으며, 학습시간 측면은 Light GBM, XGBoost 알고리즘이 우수한 성능을 보였다.

두번째로, 상용망 PCAP 파일 기반 데이터 세트의 성능평가 결과, 모든 알고리즘이 99% 이상의 높은 정확도, 정밀도, F1 Score를 보였고, 재현율은 모든 알고리즘이 Flow 기반 데이터 세트 보다 높게 나타났다(평균 0.9%). 학습시간 측면은 XGBoost, Light GBM, Random Forest 알고리즘이 상대적으로 빠른 속도를 기록하였다.

이러한 연구결과를 바탕으로 상용망 환경에서 성능과 학습시간을 종합적으로 고려할 때 XGBoost, Light GBM, Random Forest가 DDoS 탐지에 가장 적합한 알고리즘으로 평가된다.

그림 2는 상용망 2가지 유형의 데이터 세트에 대한 성능평가 결과이고, 표 2는 학습시간을 나타낸다.



(그림 2) 상용망 Flow, PCAP 데이터세트 성능평가 결과

구분	Flow Time(s)	PCAP Time(s)
XGBoost	55.10	0.59
Light GBM	90.18	0.18
랜덤포레스트	775.76	0.77
CatBoost	2434.00	3.94
GBM	1287.03	4.28

<표 2> 상용망 Flow, PCAP 데이터세트 학습시간 평가 결과

5.4 데이터 세트의 피처 중요도 분석

CIC-DDoS2019 데이터 세트와 상용망 데이터 세트에서 머신러닝 알고리즘이 실제 사용한 피처수 분석결과는 다음과 같다. CIC-DDoS2019 Flow 기반 데이터 세트는 77개 중 65개(84%)의 피처를 사용하였고, 상용망 데이터 세트는 Flow 기반에서 77개중 44개(57%), PCAP 파일기반에서 71개 중 53개(74%)의 피처를 사용하였다.

피처 중요도는 알고리즘별 계산방식, 구조, 상관관계 등에 따라 다르지만 실제 사용된 주요 피처 예시는 다음과 같다.

- Fwd Packet Length Max: 송신측 전송패킷 중 최대길이
- ttl_same_ratio: 동일한 TTL을 가진 패킷 비율
- Fwd Header Length : 송신측 전송 패킷의 헤더길이
- Total packets : 송수신된 전체 패킷의 수

특히, 상용망 데이터 세트는 실험망 데이터 세트 대비 상대적으로 적은 수의 피처를 사용하면서도 높은 정확도를 달성하였다는 점이 흥미롭다. 이는 상용망 데이터 세트에서 선택된 피처들이 DDoS 공격 탐지에 있어 더욱 효율적이고 중요한 정보를 제공할 수 있음을 시사한다. 반면, CIC-DDoS2019 데이터

세트의 일부 피처들은 중복되거나, 모델 성능 향상에 기여하지 않는 불필요한 정보를 포함하고 있을 가능성이 있다.

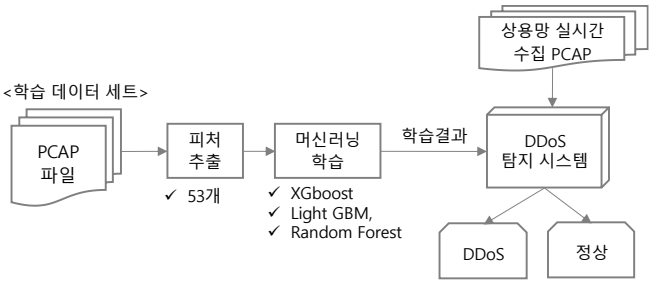
따라서, 효과적 피처 선택을 통해 모델 복잡성을 줄이면서도 높은 성능을 유지할 수 있다.

6. 상용망에 적합한 머신러닝 기반 DDoS 탐지 시스템

6.1 머신러닝 기반 DDoS 탐지 시스템

본 연구에서 제안하는 머신러닝 기반 DDoS 탐지 시스템 아키텍처는 다음과 같이 구성된다. 먼저, 학습 데이터로 사용될 PCAP 파일에서 머신러닝 알고리즘에 적합한 피처를 추출하고 전처리 한 후 머신러닝 모델을 학습시킨다. 학습된 모델은 상용망에서 실시간 수집된 PCAP 파일을 입력받아 DDoS 공격을 탐지하고, 정상트래픽과 비정상트래픽을 실시간 판별한다.

그림 3는 이러한 머신러닝 기반 DDoS 탐지 시스템 아키텍처를 시각적으로 표현한 것이다.



(그림 3) 머신러닝 기반 DDoS 탐지시스템 아키텍처

6.2 DDoS 탐지 대상 데이터 선정

상용망 환경에서 모든 머신러닝 알고리즘이 Flow 기반과 PCAP 파일기반 데이터 세트에서 99% 이상의 정확도, 정밀도, F1 Score를 보였으나 PCAP 파일 기반에서 재현율이 평균 0.9% 높고, 학습시간 측면도 PCAP 파일기반 데이터가 더 빠른 속도를 기록 하였다. 이러한 결과는 대규모 네트워크 운용환경에서 PCAP 파일 기반 DDoS 탐지가 적합한 것으로 판단된다.

6.3 최적 알고리즘 선정

상용망 환경에서 최적의 DDoS 탐지 알고리즘을 성능과 학습 시간 측면에서 종합적으로 고려한 결과, PCAP 파일 기반에서 XGBoost, Light GBM, Random Forest가 적합한 알고리즘으로 판단된다. 이들 알고리즘은 높은 정밀도와 재현율을 기록하여 실시간 탐지에서 우수한 성능을 발휘하고, 학습시간 측면에서도 상대적으로 짧은 시간을 보였다.

6.4 최적 피처 항목 선정

상용망 PCAP 파일 기반 알고리즘에서 71개 중 53개 피처를 사용하였고, 알고리즘별로는 XGBoost 34개, Light GBM 40개, Random Forest 53개로 확인되었다.

DDoS 탐지 모델의 정확도와 처리속도를 개선하기 위해 최적 피처 선정이 중요한 것으로 사료된다.

7. 결론

본 연구에서는 상용망 환경에서 DDoS 공격을 실시간 탐지하기 위한 최적의 데이터와 알고리즘을 선택하고 DDoS 탐지 시스템을 제안하였다.

실험결과, PCAP 파일기반 데이터에서 XGBoost, Light GBM, Random Forest 알고리즘이 99%의 높은 정확도와 신뢰성, 짧은 처리시간을 보여 상용망 환경에서 DDoS 탐지에 적합한 것으로 분석되었다.

향후 연구에서는 상용망의 다양한 조건과 상황에 맞춘 추가적인 피처 설계, 최적화 및 Payload를 분석하는 딥러닝 알고리즘을 추가해 DDoS 탐지 시스템을 더욱 발전 시킬 수 있을 것이다.

참고문헌

[1] 조창섭. 사이버공격 탐지성능개선을 위한 머신 러닝기반 보안관제시스템. 숭실대학교 박사논문, 2019년 6월

[2] Sharafaldin, A. H. Lashkari, S. Hakak and A. A. Ghorbani, "Developing Realistic Distributed Denial of Service (DDoS) Attack Dataset and Taxonomy," 2019 International Carnahan Conference on Security Technology (ICCST), Chennai, India 2019, pp. 1-8, doi: 10.1109/CCST.2019.8888419.

[3] T. Subbulakshmi, K. BalaKrishnan, S. M. Shalinie, D. AnandKumar, V. GanapathiSubramanian and K. Kannathal, "Detection of DDoS attacks using Enhanced Support Vector Machines with real time generated dataset," 2011 Third International Conference on Advanced Computing, Chennai, India, 2011, pp. 17-22, doi: 10.1109/ICoAC.2011.6165212.

[4] M. Al-Eryani, E. Hossny and F. A. Omara, "Efficient Machine Learning Algorithms for DDoS Attack Detection," 2024 6th International Conference on Computing and Informatics (ICCI), New Cairo - Cairo, Egypt, 2024, pp. 174-181, doi: 10.1109/ICCI61671.2024.10485168.

[5] J. Buzzio-García et al., "Exploring Traffic Patterns Through Network Programmability: Introducing SDNFlow, a Comprehensive OpenFlow-Based Statistics Dataset for Attack Detection," in IEEE Access, vol. 12, pp. 42163-42180, 2024, doi: 10.1109/ACCESS.2024.3378271.

[6] 서인혁, 이기택, 유진현, & 김승주. (2017). CNN 기반의 실시간 DNS DDoS 공격 탐지 시스템. 정보처리학회논문지. 컴퓨터 및 통신시스템, 6(3), 135-142