Entropy 기반 DDoS 탐지 회피공격

이지원, 김경백
전남대학교 정보보안협동과정

Entropy-based DDoS Detection Evasion Attack

Jiwon Lee, Kyungbaek Kim
Interdisciplinary program of Information Security,
Chonnam National University

요 약

급격히 증가하는 IoT 장비들을 통해 사람들의 편의성은 증가하였다. 반면 Mirai 봇넷과 같이, 공격자들은 IoT를 감염시켜 공격수단으로 사용하고 있다. 이러한 문제의 해결을 위해, IoT 장비에서 발생하는 네트워크 트래픽을 기계 학습을 이용하여 공격을 탐지하는 연구가 활발하게 진행되었다. 하지만 대표적인 DDoS 탐지 방법 중 하나인 entropy 수치를 기반의 탐지 모델의 경우, 공격자는 다수의 감염된 IoT 장비들을 사용해 entropy 수치를 조작하여 공격 탐지 회피가 가능하게 되었다. 이 논문에서는 이러한 entropy 수치 조작을 통한 공격 탐지 회피의 가능여부를 확인한다. 이를 위해, entropy 기반 DDoS탐지 기계학습 모델을 구현하였고, entropy 수치 조작을 통해 DDoS 탐지 회피가 가능하다는 것을 확인하였다.

I. 서론

현재 다양한 DDoS 탐지 연구들이 있으며, 대표적인 연구로는 entropy 수치를 기반으로 하는 탐지 모델이 있다. 시각적으로도 수치를 확인할 수 있으므로 현재까지도 많은 연구가 이루어지는 중이다[1]. 그 중에서 entropy를 기반으로 하는 DDoS 탐지 기술은 아직도 여러 범위에서 사용되고 있다[2]. 하지만 ip 주소 값을 entropy 수치를 탐지로 사용할 경우, 공격자가 IoT 장비를 감염시켜 봇으로 만들어 이 entropy기반 DDoS 탐지를 회피할 가능성이 있다.

이러한 공격의 가능성을 확인하기를 위하여 우리는 CIC-IDS2017 데이터 셋의 ip와 port에 대한 엔트로피 수치를 이용한 모델을 구현하였다[3]. 그 후 가상의 패킷을 생성하여 엔트로피 수치를 제어하는 시나리오를 기반으로 entropy 수치를 조작하여 공격 탐지 회피가 가능한지 실험하였다

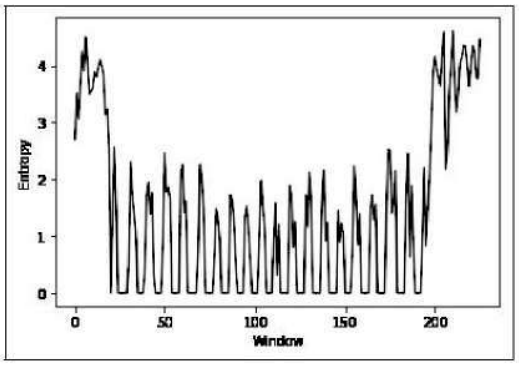
II. 관련연구

현재 데이터의 증가와 하드웨어의 발달로 기계 학습 모델의 연구 활발하나 정제된 데이터 셋을 구하기에는 아직 부족하다. 이 때문에 필요한 트래픽 데이터를 생성하는 연구가 이루어지고 있다. 일부 연구에는 부족한 데이터 셋의 볼륨이나 데이터 셋의 불균형을 해결하기 오버 샘플링 기법을 사용하여 모델 성능을 올리고자 연구도 있다[4]. 다른 연구로는 seqGAN를 활용한 트래픽 생성으로 탐지 모델 성능을 개선에 중점 연구가 있는 반면에[5], R Chauhan는 WGAN을 이용해 주요 피처를 하나씩 추가로 학습하여 감지하기 힘든 악성 트래픽을 만드는 데 중점을 두었다[6]. 우리는 entropy 기반의 DDoS 탐지 회피가 가능한 악성 트래픽을 제작하여, 다수의 감염 봇을 확보한 공격자의 탐지 회피 위협 정도를 평가하였다.

Ⅲ. Entropy 기반 DDoS 탐지 회피 공격

$$H(X) = -\sum p_x \log p_x \quad (1)$$

Entropy는 무질서함을 수치로 나타낸 것으로, 식 1과 같이 정의된다. 데이터 셋을 패킷 수에 따라 분할하여 각 윈도우 들의 entropy를 계산하여, 수치를 비교한다.



[그림 1] Source IP에 대한 entropy 수치

그림 1은 CIC-IDS2017 데이터 셋의 'source ip'대하여 계산한 entropy 수치이며, 윈도우 30에서 DDoS 공격이 일어나자 entropy 값이 급격히 내려가는 것을 볼 수 있다. 이처럼 공격에 의한 네트워크의 급격한 변화를 포착하는데 유용하게 쓰인다. 이 논문에서는 이러한 entropy 수치를 ip와 port값을 조작하여 급격한 entropy 수치를 제거 했을 때, 이를 이용한 탐지 모델의 성능을 비교하고자 한다.

3. 1 탐지 모델 생성

우리는 데이터 셋에 'source ip', 'destination ip', 'source port', 'destination port'의 entropy 수치를 매 1000개의 패킷의 발생점을 기준으로 ±1000 A그룹, 매 100개의 패킷의 발생점을 기준으로 ±100을 계산 후 생성한 B그룹, 추가로 'Protocol'와 'Fwd Packets/s', 'Bwd Packets/s'을 3가지 피처를 C그룹으로 만들어 진행하였다. 이후 데이터 셋을 7:3 비율로 나눈 뒤에 K-최근접 이웃(KNN), Support Vector Machine(SVM), Logistic Regression(LR), Deep Neural Network(DNN) 모델을 만들어 그룹을 조합하여 학습 및 테스트하였으며 결과는 표 1과 같다.

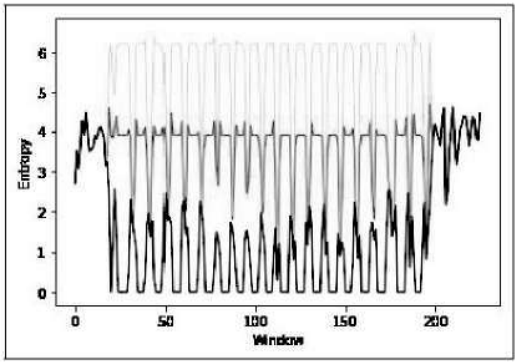
표 1을 보면 entropy 수치만을 이용한 A, B 그룹도 높은 정확도를 보여주고 있다. 또한 window 범위를 넓게 측정한 A그룹이 B그룹보다 좋은 성능을 나타내고 있다.

[표 1] 기계학습별 정확도

모델	A	B	C	A+C	B+C
KNN	96.51	98.01	80.67	99.46	99.63
SVM	95.07	79.72	68.51	95.19	98.77
LR	94.83	79.63	68.78	95.28	99.96
DNN	94.83	71.80	71.44	95.67	95.67

3.2 회피 DDoS 공격 트래픽 제작

준비된 데이터 셋에는 일반 DDoS 공격이 대부분이지만, 애플리케이션 층을 이용한 slowloris 혹은 hulk 공격 등이 존재하며, 이러한 Port 번호 변경이 불가능한 애플리케이션 계층 DoS 공격을 제외한 포트를 변경하기로 하였다. 또한 Destination IP는 공격 지점이기에 변경 대상이 아니지만, Source IP는 공격자가 봇을 활용하여 공격 가능한 IP 수를 조절할 수 있다. 위와 같은 정보와 공격 대상의 네트워크 대역에 스니핑이 가능하다는 시나리오를 중심으로 실험을 진행했다.



[그림 2] 공격 IP 대역 할당의 따른 entropy 수치 변화

Entropy 수치 계산 방법은 매 1000개의 패킷의 발생점을 기준으로 ±1000단위로 고정하였고, 정상 트래픽에서 발생하는 시간별 Port 번호 개수와 IP 개수를 참고해 대입하였다. 그림 2에서 빨간색 선은 DoS에서 사용하는 IP 주소를 15배로, 노란색은 75배 비율로 늘려서 사용하였으며, 15배일 경우 정상 트래픽과 유사하게 entropy 값을 보여주고 있다. Source Port도 이처럼 유사한 값을 출력해주었지만, Destination Port는 애플리케이션 계층 DoS를 사용하기에 이전 entropy 수치랑 비교해 큰 변화가 없었다.

우리는 이전에 생성했던 기계학습 모델로 생성한 데이터를 실험하였으며, 피처는 앞서 말했던 4가지 entropy 수치와 'Protocol'와 'Fwd

Packets/s', 'Bwd Packets/s'를 가지고 실험했다. 표 2에서 entropy 수치를 중점으로 한 기계학습 모델들의 정확도 하락을 확인할 수 있었다.

[표 2] 생성한 모델을 사용했을 경우 기존 모델의 정확도

모델	15배	75배
KNN	42.98	94.54
SVM	99.86	91.23
LR	96.69	42.86
DNN	85.66	88.54

IV. 결론

이 논문에서는 다수의 감염된 봇을 확보한 공격자가 entropy 수치 기반의 공격 탐지를 회피 할 수 있는 가능성을 제시한다. 실험 결과 entropy 수치 중심의 공격 탐지 모델의 경우 다수의 감염된 IoT 장비를 확보한 공격은 큰 위협이 될 수 있는 것을 확인하였다. 대부분 기계학습 모델에서는 이보다 많은 피처를 가지고 사용하나, Feature Selection같이 중요도 높은 피처를 선별하여 만들거나, DDoS 탐지의 주요 지표인 entropy 중심의 모델의 경우에는 이러한 탐지 회피 공격에 안심할 수 없다. 이번 실험의 한계점은 데이터 셋을 이용하여 공격 트래픽을 생성하는 방식이었다. 향후 연구에는 가상 네트워크 구성하고 실측 네트워크 트래픽을 활용하여 entropy 수치에 따른 공격 가능성을 실험해 볼 예정이다. 또한, GAN등을 통한 인공지능 기반 공격 탐지 회피용 공격 트래픽 생성기술에 대한 연구도 추진할 계획이다.

ACKNOWLEGEMENTS

"이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임"(IITP-2019-0-01343)

[참고문헌]

[1] K. Kalkan, L. Altay, G. Gür and F. Alagöz, "JESS: Joint Entropy-Based DDoS Defense Scheme in SDN," in IEEE Journal on Selected Areas in Communications, vol. 36, no. 10, pp. 2358-2372, Oct. 2018, doi: 10.1109/JSAC.2018.2869997.

[2] R. Wang, Z. Jia and L. Ju, "An Entropy-Based Distributed DDoS Detection Mechanism in Software-Defined Networking," 2015 IEEE Trustcom/BigDataSE/ISPA, 2015, pp. 310-317, doi: 10.1109/Trustcom.2015.389.

[3] Iman Sharafaldin, Arash Habibi Lashkari, and Ali A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization", 4th International Conference on Information Systems Security and Privacy (ICISSP), Portugal, January, 2018

[4] 이종화, 방지원, 김종욱 and 최미정. 데이터의 불균형성을 제거한 네트워크 침입 탐지 모델 비교 분석. KNOM Review, 23(2), 18-28. 2020

[5] 이우호, 함재균, 정현미 and 정기문, 네트워크 공격 탐지 성능향상을 위한 딥러닝을 이용한 트래픽 데이터 생성 연구. 한국융합학회논문지, 10(11), 1-7. 2019

[6] R. Chauhan and S. Shah Heydari, "Polymorphic Adversarial DDoS attack on IDS using GAN," 2020 International Symposium on Networks, Computers and Communications (ISNCC), 2020, pp. 1-6, doi: 10.1109/ISNCC49221.2020.9297264.