IoT 보안 위협에 대한 딥러닝 기반의 탐지 동향

김현지*, 서화정**†
*한성대학교 IT 융합공학부
**† 한성대학교 IT 융합공학부
khj1594012@gmail.com, Hwajeong84@gmail.com

Trends in detection based on deep learning for IoT security threats

Hyun-Ji Kim*, Hwa-Jeong Seo**†
*Dept. of IT Convergence Engineering, Han-sung University
**† Dept. of IT Convergence Engineering, Han-sung University

요 약

최근 5G, 인공지능(AI) 등과 함께 사물인터넷 (IoT) 기술이 주목받고 있으며, 보안 위협 또한 증가하고 있다. IoT 기기에 대한 다양한 공격 기법들이 존재하는 만큼 IoT 보안에 관한 연구 또한 활발하게 진행되고 있다. 본 논문에서는 IoT 환경에서의 보안 위협에 대응하기 위한 딥러닝 기반의 탐지 기법들의 최신 연구 동향과 앞으로의 방향을 살펴본다.

1. 서론

최근 AI, 엣지 컴퓨팅, 5G 등 ICT 기술이 발전하면서 사물인터넷 (IoT) 분야가 더욱 주목받고 있으며, 그 규모가 증가함에 따라 보안 위협 또한 증가하였다. 사물인터넷 기기들이 연결된 네트워크, 어플리케이션 등 다양한 측면에서의 공격들이 존재하며 [1], IoT 보안의 중요성이 더욱 부각되고 있다.

2. 관련 연구

2.1 딥러닝 (Deep Learning)

딥러닝이란 인공 신경망을 기반으로 하는 기계 학습의 일부이다. 입력층과 출력층을 포함한 여러 계층들을 사용하여 입력 데이터로부터 더 고차원적인 특징을 추출하여 학습한다. 각 뉴런이 갖는 가중치들을 더 나은 방향으로 가중치를 학습하여 분류 및 회귀 등의 문제를 해결하며, 가중치 학습은 손실함수의 값을 반영하여 조절된다. 다양한 구조의 딥러닝 모델들이 있으며, 해결해야 할 문제에 맞는 파라미터와 모델을 사용해야한다.

딥러닝 모델 종류로는 이미지 데이터 학습에 강점을 갖는 합성곱 신경망 (Convolutional Neural Network), 시점이나 순서가 있는 시퀀스 데이터에 주로 사용하는 순환 신경망 (Recurrent Neural Network), 노드와 엣지를 갖는 그래프 데이터에 사용하는 그래프 신경망 (Graph Neural Network), 데이터를 생성해내는 생성형 적대 신경망 (Generative Adversarial Network)등이 있으며, 각 구조들에 대해서도 많은 변형들이 존재한다.

또한, 평가 방법은 주로 정밀도와 재현율의 조화평균인 F-measure를 사용하며, 1에 가까울수록 좋은 성능을 나타낸다.

2.2 인공지능과 사물인터넷

최근 인공지능이 사물인터넷과 융합되어 AIoT (AI + IoT) 라는 분야로 발전하고 있다. 사물 인터넷에 연결된 기기의 수, 고속 및 고 대역폭 네트워크로 인한 다양하고 거대한 데이터들을 처리하기 위해 인공지능과 사물인터넷이 결합된 형태이다.

또한 데이터 처리뿐만 아니라 사물인터넷의 보안성을 향상시키기 위한 연구도 진행되고 있다.

3. 딥러닝 기반의 IoT 보안 위협 탐지 동향

3.1 노드 및 게이트웨이 공격 [2]

IoT에 사용되는 엣지 노드 및 게이트 웨이에 대한 공격을 탐지하여 민감한 트래픽은 다시 라우팅하여 다른 경로로 전송하는 기법이다. 즉, 자동적으로 엣지 노드의 보안 위협을 감지하며 서비스의 품질과 효율성을 위한 실시간 대응이 가능해진다.

모바일 장치, 게이트 웨이, 서버 등이 연결된 그래

프 형태이다. 따라서 그래프 신경망으로 구성되며 학습을 통해 IoT 네트워크에 대한 공격을 탐지할 수 있다.

3.2 ARP spoofing[3]

ARP 프로토콜의 허점을 이용하여 자신의 MAC 주소를 다른 컴퓨터의 주소인 것처럼 속이는 공격이다. 따라서, 세션 하이재킹 또는 중간자 공격을 가능하게 한다. 이는 IoT 기기에 있어 치명적인 위협이 될 수 있으며 딥러닝 기반의 스푸핑 공격 탐지에 관한 연구 또한 진행되었다. 데이터 셋은 TCP, UDP, 그리고 ARP 패킷들의 1분당 백분율으로 구성되며, 신경망을 통해 훈련된 후, 스푸핑 공격을 탐지한다.

3.3 스펙터 (spectre) 및 멜트다운 (Meltdown)[4]

스펙터와 멜트다운은 각각 실행시간 최적화를 위한 추측 실행과 비순차 실행의 과정에서 발생하는 취약점을 이용한 공격이다. 해당 취약점들은 운영체제의 메모리 보호 기능을 우회하여 비밀 데이터를 스니핑 할 수 있다. 운영체제 패치, 구조적 변경, 동적 분석 등으로 대응하고 있으며, 다양한 딥러닝 모델을 활용한 탐지 기법들이 연구되고 있다.

생성형 딥러닝 네트워크를 통해 생성한 가젯과 실제 샘플을 활용하여 학습한 후, 15개의 스펙터 샘플을 100만 개의 가젯으로 확장한 후, 생성형 적대 신경망을 통해 해당 가젯들의 분포를 학습한다. 즉, 조건부 GAN인 MASK GAN을 사용하며, 이를 통해 새로운 스펙터 가젯을 생성한다.

MASK GAN은 어텐션 기법이 있는 seq2seq 구조를 기반으로 하는 조건부 GAN의 한 종류이다. 또한 어텐션 메커니즘을 통해 각 입력 시퀀스들은 출력에 대해 서로 다른 가중치를 가질 수 있다. MASK GAN의 가장 큰 특징은 주변 데이터들의 컨텍스트에 따라 일부 토큰이 마스킹된다. 따라서 더 길고 유의미한 시퀀스를 생성할 수 있는 모델이다.

가젯 생성 이후, 자연어 처리 모델 중 하나인 BERT를 기반으로 한 고차원 단어 임베딩을 사용하여 스펙터 가젯을 탐지한다. 해당 기법은 0.99의 F-measure를 달성하며 높은 성능을 보였다.

3.4 멀웨어(Malware) 탐지[5]

IoT환경에서 인공지능을 통한 멀웨어 탐지 방법이 다양하게 연구되었다. 네트워크 상의 패킷을 대상으로 하거나, 정적 분석을 통해 멀웨어를 탐지한다. 특히, CNN(Convolutional Neural Network)를 기반으로 하여 멀웨어를 탐지해내는 연구가 다수 진행되었다.

3.4.1 제어 흐름 그래프(control flow graph, CFG) 기반의 멀웨어 탐지[5]

노드 수, 엣지 수 등을 분석할 수 있는 제어 흐름 그래프 (control flow graph, CFG)를 사용하여 멀웨어 바이너리에 대한 분석을 수행한다. 즉, 두 멀웨어 유형의 CFG를 비교하여 바이너리를 대조하여 분류한다. 학습을 위해 CFG를 기반으로 추출된 그래프 특징은 최단거리, 희소성, 엣지의 수, 노드의 수 등이 있다. 비정상적 실행 흐름이 발생할 경우 해당 특징들이 갖는 값들의 특성이 달라지므로 정상 작동과 멀웨어로 인한 비정상 작동을 분류할 수 있게 된다.

또한 해당 연구에서는 적대적 공격 방법인 Graph Embedding and Augmetation (GEA)를 사용한다. 적대적 공격 방법이란 정상적인 학습 데이터에 적대적 데이터를 추가하여 두 데이터를 가려내는 성능을 향상시키는 것이다. 즉, 해당 모델의 일반화 성능을 높이는 것이다. GEA 방식을 통해 원본 샘플의 특성을 보존하며, 100%의 높은 오분류율을 달성한다.

그래프를 학습 데이터로 사용하는 그래프 신경망을 통해 제어 흐름 그래프의 특징을 이미지로 바꾸지 않고서도 학습이 가능할 것으로 보인다.

3.4.2 부채널 분석을 활용한 멀웨어 탐지

인공지능 기반의 다양한 부채널 공격 탐지 기법이 있다. 딥러닝 기반의 부채널 분석 공격을 통해 IoT 디바이스를 타겟으로 하는 멀웨어를 탐지한다.

IoT 디바이스의 원시 전력 파형을 필터링하여 랜섬웨어에 해당하는 파형을 얻어낸다. 랜섬웨어 파형을 학습한 후, 해당 디바이스에서 지속적으로 전력 파형들을 수집하여 감염 여부를 확인한다.

학습에 사용되는 모델은 하나의 시퀀스를 다른 시퀀스로 바꾸는 두개의 RNN이 함께 동작하는 seq2seq 모델이다. 입력 데이터로 사용되는 전력 파

형은 시계열 데이터이므로 RNN을 사용하였으며, 어텐션 메커니즘을 함께 사용하여 입력 시퀀스의 특정 범위에 가중치를 두어 더 집중할 수 있도록 한다. 해당 연구는 5개의 멀웨어에 대해 평균 90.36%의 정확도를 달성하였다.

이 외에도 IoT 디바이스에서 사용되는 센서 데이터를 기반으로 멀웨어를 탐지하는 연구 또한 진행되었다. 바이너리 머신 상태를 구분하는 특징 벡터를 형성하여 로지스틱 회귀 접근 방식을 통해 기계학습을 수행한다. 로지스틱 회귀를 이진분류기로 사용하여 생성된 확률 추정치를 임계 값과 비교하여 해당 예측 상태가 거짓일 경우 정상 작동이 되고, 아닌 경우 랜섬웨어 실행으로 분류되어 IoT 환경에서 랜섬웨어를 탐지하고 예방할 수 있다.

3.5 봇넷(botnet) 탐지

봇넷은 멀웨어의 한 종류인 봇이라는 소프트웨어에 의해 감염된 다수의 좀비 컴퓨터로 구성되어있는 네트워크이다. DDoS 공격, 개인정보 수집, 스팸 메일 전송 등에 이용되며, 2016년 최초로 사물인터넷을 이용한 봇넷이 등장하였다. 임베디드 기기에 접속하여 악성코드를 전파시키는 방법으로 봇넷을 구성하고, 이를 이용하여 DDoS 공격을 수행한다. 미라이 봇넷의 경우 보안이 허술한 IoT기기에 해당 악성코드가 배포되어 트래픽을 라우팅하는 DYN 서버를 공격한다. 봇넷 공격은 1테라 비트의 초당 비트수를 가지며, 이는 기존의 DDoS 공격에 비해 큰 수치이다. 이는 사물인터넷 환경에서 큰 보안 위협이 되는 공격이며, 딥러닝을 활용하여 봇넷을 탐지하고 예방하기 위한 연구가 진행되었다.

3.5.1 CNN을 활용한 봇넷 탐지[6]

일반적인 트래픽과 봇넷 트래픽을 캡처하고, 패킷 평균 길이, 세션 수, UDP에 대한 TCP 패킷의 비율, DNS 개수 등의 특징을 추출한다. DNS 개수는 일반 트래픽의 경우 UDP 패킷 수에 비해 굉장히 적었으며, 봇넷의 경우는 DNS 패킷 수가 UDP 패킷에 대해 높은 비율을 차지하였다. 플로우 개수는 봇넷의 경우 호스트와 짧은 간격으로 통신함을 나타냈다. 이러한 특징점들을 찾아내어 전처리한 데이터들은 그림 1과 같이 이미지로 변환되어 학습에 사용된다. 즉, 각 트래픽들의 특징들이 각각 일반, 봇넷 트래픽으로 분류되는 형태이다. CNN을 활용하여 분류한 결과 일반 트래픽의 경우 100%로 분류되었고, 봇넷 트래픽의 경우 약 98.7%는 정확하게 분류하였고 오탐률은 1.3%를 달성하였다.

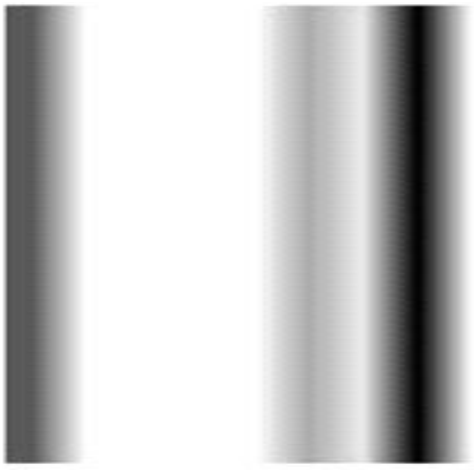

그림 1. 봇넷 패킷 특징값 시각화 예시

3.5.2 BLSTM-RNN 을 활용한 봇넷 탐지[7]

Bidirectional LSTM-RNN (BLSTM-RNN)의 경우 과거의 데이터 뿐만 아니라 미래의 데이터까지 반영하여 학습 가능한 네트워크이다. 사용된 모델은 sigmoid 활성화 함수, Mean Absolute Error 손실 함수, Adam 최적화 함수를 사용하였으며, BLSTM 레이어는 총 6개의 유닛을 가진다. 각 패킷들을 입력 데이터로하여 그에 해당하는 공격으로 분류하는 네트워크이며 100번의 학습을 진행하였다. 미라이 봇넷, UDP, DNS 공격 유형에 대해서는 각각 99.9%, 98.5%, 98.4%로 높은 정확도를 달성하였으나, 가장 샘플 수가 많았던 ACK 공격에 대해서는 93.7%로 가장 낮은 정확도를 보였다. 또한 ACK를 포함한 멀티벡터의 경우는 포함하지 않은 경우보다 5.6% 더 낮은 결과를 보였다. 이는 ACK 패킷의 시퀀스 번호 등의 특성과 복잡성 때문인 것으로 추측된다.

4. 결론

최근 주목 받고 있는 사물 인터넷과 인공 지능 기술을 결합하여, IoT 환경에서의 보안 위협을 방어하기 위한 방법들이 연구되고 있다. 여러 측면에서의 보안 위협에 대응할 수 있도록 다양한 딥러닝 모델들을 적용해보고, 보안 위협의 탐지를 넘어서 즉각적인 대응이 가능하도록 하여 IoT 서비스의 품질 및 안전성을 유지할 수 있는 방법들이 연구되어야 할 것으로 생각된다.

참고문헌

[1] I. Butun, P. Österberg and H. Song, "Security of the Internet of Things: Vulnerabilities, Attacks, and Countermeasures," in IEEE Communications Surveys & Tutorials, vol. 22, no. 1, pp. 616-644, Firstquarter 2020.

[2] E. Gelenbe et al., "IoT Network Attack Detection and Mitigation," 2020 9th Mediterranean Conference on Embedded Computing (MECO), Budva, Montenegro, pp. 1-6, 2020

[3] Abdulla, Husain and Al-Raweshidy, Hamed and S. Awad, Wasan, "ARP Spoofing Detection for IoT Networks Using Neural Networks," February 10, 2020.

[4] Tol, M.C., Yurtseven, K., Gülmezoglu, B., & Sunar, B. "FastSpec: Scalable Generation and Detection of Spectre Gadgets Using Neural Embeddings,". ArXiv, abs/2006.14147, 2020.

[5] A. Abusnaina, A. Khormali, H. Alasmary, J. Park, A. Anwar and A. Mohaisen, "Adversarial Learning Attacks on Graph-based IoT Malware Detection Systems," *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*, Dallas, TX, USA, pp. 1296-1305, 2019.

[6]Lee-SangHyeok,."BotNet Detection Based on Deep Learning." master's thesis, Sungkyunkwan University, 2017.

[7] C. D. McDermott, F. Majdani and A. V. Petrovski, "Botnet Detection in the Internet of Things using Deep Learning Approaches," 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, pp. 1-8. 2018.