사물 인터넷 환경에서 경량화된 침입 탐지 기술에 대한 연구

송정환*, 권태경
서울대학교 컴퓨터공학부

*jhsong@mmlab.snu.ac.kr, tkkwon@snu.ac.kr

A Study on the light-weight Intrusion Detection System (IDS) in the Internet of Things (IoT)

Junghwan Song*, Ted “Taekyoung” Kwon
Seoul National Univ.

요 약

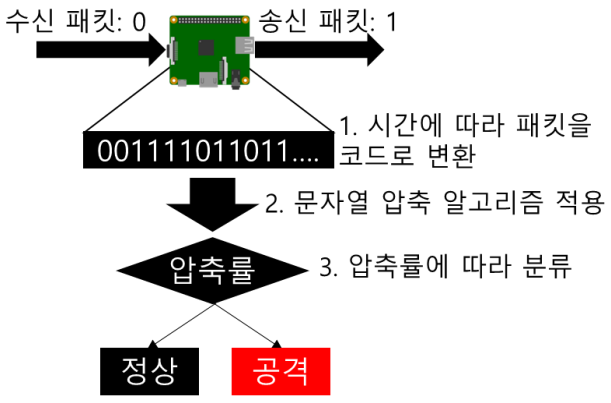
본 논문은 최근 사물 인터넷 기기들이 급속도로 증가함에 따라 그 위험성도 함께 급증한 사물 인터넷에서의 네트워크 위협을 방지하기 위한 침입 탐지 기술에 대해 연구하였다. 특히, 사물 인터넷 기기에서 동작할 수 있는 경량화된 침입 방지 시스템을 제안하는데 중점을 두었다. 이러한 경량화된 침입 방지 시스템의 실험을 통해 성능을 검증하였다.

Ⅰ. 서 론

최근 사물 인터넷의 매우 가파른 성장세가 주목받고 있다. 이러한 흐름에 따라 사물 인터넷에 연결된 단말들의 숫자가 폭발적으로 증가하고 있으며, 이는 사물 인터넷으로 인한 새롭고 다양한 네트워크 위협을 초래하였다.[1] 대표적인 예로, 사물 인터넷 기기가 봇넷에 감염되어 DDoS 공격에 사용되는 경우를 들 수 있다. 기존의 데스크톱은 물론이고, 스마트폰으로 이어지는 모바일 기기보다 사물 인터넷 기기들은 그 수에서 양적으로 수십 배 이상 크기 때문에 이 기기들이 감염되어 발생시키는 트래픽 역시 매우 크고, 이는 효과적인 DDoS 공격이 가능하게 한다. 이러한 사물 인터넷의 네트워크 위협점을 방지하기 위해 다양한 침입 탐지 시스템들이 제안되었다. 그러나 대부분의 제안된 침입 탐지 시스템은 프록시 서버 등 네트워크 트래픽이 모이는 지점에서 침입자를 분류해내는 방식이다.[2] 따라서 상대적으로 여유로운 컴퓨팅 자원 (연산 속도, 메모리 크기 등)이 확보되어 있기에 주로 기계 학습을 활용해왔다. 그러나 사물 인터넷에서 내부 네트워크 트래픽이 집중되는 지점이 컴퓨팅 자원이 충분하지 않을 경우도 있으며, 사물 인터넷 내부의 프록시 서버나 게이트웨이를 거치지 않고 바로 외부의 서버로 데이터를 전송하는 형태도 존재할 수 있다. 따라서 본 논문에서는 게이트웨이나 프록시가 아닌 사물 인터넷 기기 자체에서 동작하기 위한 경량화된 침입 탐지 시스템을 제안하고자 한다.

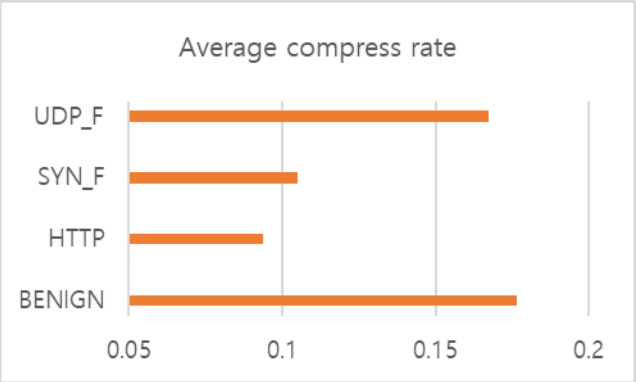
Ⅱ. 본론

침임 탐지 시스템을 경량화하기 위해 본 논문에서는 기계 학습을 사용하지 않는 침임 탐지 시스템을 제안한다. 이를 위해 [그림 1]과 같이 네트워크 트래픽을 수신, 송신 패킷으로 나누어 각각 1 비트의 (1, 0)으로 치환하고, 치환된 데이터에 문자열 압축 알고리즘을 적용하여 압축률에 따라 공격 여부를 판단하는 침입 탐지 알고리즘을 고안하였다. 이 방식은 공격 트래픽의 경우 같은 패턴이 반복되는 점, 특히 DoS 나 DDoS 등의 공격의 경우 사물 인터넷 기기가 감염되어 공격자가 되었을 경우 송신 패킷의 숫자가 상대적으로 높은 점, 공격 대상이 되었을 반대의 현상이 일어나는 점을 활용한 것이다.



[그림 1. 경량화된 침입 방지 시스템의 동작 방식]

이 때 [그림 1]의 2 번에서 어떤 문자열 압축 알고리즘을 사용하는가에 따라 시스템의 성능이 크게 변화한다. 본 논문에서는 Adaptive Huffman coding[4] 방식을 사용하여 문자열을 압축하였다. [4] 알고리즘은 실시간으로 추가되는 문자열에 대해 치환을 행하면서 압축하기에 적합하다. 해당 알고리즘을 사용한 침입 방지 시스템의 성능을 평가하기 위해 직접 테스트베드를 구성하여 3 종류의 네트워크 공격 및 정상 트래픽을 수집하였다. 3 종의 공격 트래픽은 Mirai 봇 넷의 오픈 소스를 활용하여 DDoS 공격의 일종인 UDP flooding, HTTP flooding, 그리고 SYN flooding 을 발생시켰다. 이에 따른 Huffman coding 및 Adaptive Huffman coding 의 결과는 [그림 2]와 같다.



[그림 2. Adaptive Huffman coding 을 사용한 트래픽 압축률 결과]

[그림 2]의 결과에서 압축률은 (압축 결과 길이 / 원래 문자열 길이) 를 통해 계산하였다. 따라서 값이 낮을수록 압축이 잘된 것으로 SYN/HTTP flooding 공격 트래픽들은 0.1 가량의 압축률을 보였다. 정상 트래픽은 0.175 가량의 압축률을 보임으로써 공격 트래픽과 압축률 차이가 있음을 알 수 있다. UDP flooding 의 경우 압축률이 정상 트래픽과 비슷한데, 해당 트래픽을 확인한 결과 송신 트래픽 및 수신 트래픽의 용량 차이가 크지 않아서 이러한 결과가 나온 것으로 확인하였다. 좀 더 정교한 환경에서 실험을 진행하면 정제된 결과를 얻을 수 있을 것으로 예상된다.

Ⅲ. 결론

그 동안의 사물 인터넷 환경의 침입 탐지 기술 연구는 주로 사물 인터넷의 게이트웨이에서 충분한 컴퓨팅 자원을 활용한 기계 학습 기반이었으나, 이러한 게이트웨이를 활용하기 힘든 환경을 위해 본 논문에서는 사물 인터넷 환경에서 기기에서 동작할 수 있는 경량화된 침입 탐지 기술에 대하여 연구하였다. 이를 위해 네트워크 트래픽을 비트의 연속으로 치환하고, 치환된 데이터에 대해 문자열 압축 알고리즘을 적용하여 그 압축률이 공격 트래픽과 정상 트래픽이 다름을 보였다. 후속 연구로 이를 보다 정교하게 활용하여 범용성있게 활용할 수 있는 경량화된 침입 탐지 기술을 제안하고자 한다.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구센터육성지원사업의 연구결과로 수행되었음 (IITP-2022-2020-0-01602)

참 고 문 헌

[1] Davis, Brittany D., Janelle C. Mason, and Mohd Anwar. "Vulnerability studies and security postures of IoT devices: a smart home case study." IEEE Internet of Things Journal 7.10 (2020): 10102-10110.

[2] Thakkar, Ankit, and Ritika Lohiya. "A review on machine learning and deep learning perspectives of IDS for IoT: recent updates, security issues, and challenges." Archives of Computational Methods in Engineering 28.4 (2021): 3211-3243.

[3] Vitter, Jeffrey Scott. "Design and analysis of dynamic Huffman codes." Journal of the ACM (JACM) 34.4 (1987): 825-845.