하이브리드 머신러닝 기반의 서비스거부공격과 랜섬웨어 탐지 및 대응

유다은[1], 지승하[2], 이일구[3]
[1]성신여자대학교 융합보안공학과 학부생
[2] 성신여자대학교 융합보안공학과 석사과정
[3] 성신여자대학교 융합보안공학과 부교수

yudaeun41@naver.com, 220256039@sungshin.ac.kr, iglee19@gmail.com

Hybrid Machine Learning-based Detection and Mitigation Techniques for Denial-of-Service and Ransomware Attack

Da-Eun Yu, Seung-Ha JEE, Il-Gu Lee
Dept. of Convergence Security Engineering, Sungshin Women's University

요 약

본 연구는 DDoS와 랜섬웨어 두 가지 공격을 동시에 탐지할 수 있는 하이브리드 머신러닝 모델과 탐지 이후 데이터 복구를 위한 XOR 기반 복구 기법을 제안한다. 실험 결과에 따르면 제안한 하이브리드 모델은 CNN 대비 평균 17.3%, GRU 대비 2.36% 높은 정확도를 보였으며, XOR 기법을 적용한 경우 복원율이 100%로 유지되어 효과적인 복구 성능을 입증하였다.

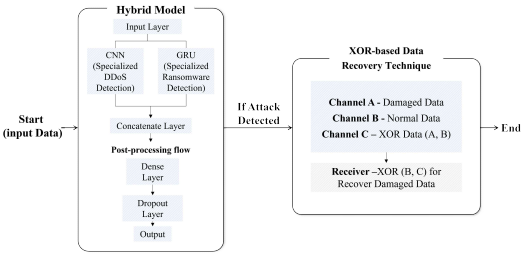
1. 서론

최근 사이버 보안 위협이 증가함에 따라, 랜섬웨어(Ransomware) 및 분산 서비스 거부(DDoS, Distributed Denial of Service) 공격 빈도가 증가하고 있다. Cloudflare (2024)의 DDoS 보고서에 따르면, DDoS 공격 중 12%가 랜섬디도스(Ransom DDoS) 공격으로 이는 2023 대비 25% 증가한 수치이며 DDoS 공격 차단 건수는 2023년 대비 53% 증가했다[1,2]. 이러한 공격 규모와 빈도가 증가함에 따라, 네트워크 공격 탐지를 위한 머신러닝 기법이 연구되고 있다.

Shubrika-Sharma[3]는 CNN(Convolutional Neural Network)과 GRU(Gated Recurrent Unit)를 융합한 하이브리드 모델이 단일 모델보다 탐지 정확도가 높음을 입증했고, Meenakshi-Mittal[4]는 딥러닝을 활용한 기법이 기존의 통계적 기법, 얕은 머신러닝(SML, Shallow Machine Learning) 기법보다 공격 탐지에서 더 효과적임을 입증하였다. 그러나 종래 연구들은 단일 머신러닝 모델을 사용한 특정 공격 탐지 최적화에만 초점을 두고 탐지 정확도와 속도를 동시에 고려하지 않았으며, 공격 대응 방법을 제안하지 않았다. 본 연구에서는 하이브리드 모델을 통한 랜섬웨어와 DDoS 공격 탐지 및 대응으로 사용자의 가용성 개선을 위한 XOR 연산 기반의 데이터 복구 기법을 제안한다.

2. 하이브리드 머신러닝 기반의 서비스거부공격과 랜섬웨어 탐지 및 대응

본 장에서는 DDoS와 랜섬웨어를 탐지하고 대응하기 위한 하이브리드 모델 기반 탐지 및 XOR 기반 데이터 복구 기법의 동작 방식을 서술한다.



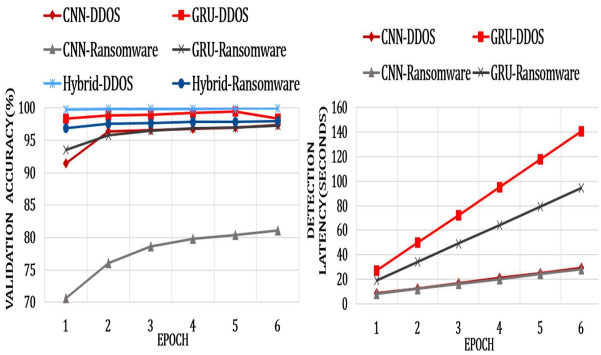
(그림 1) 하이브리드 머신러닝 기반의 서비스거부공격과 랜섬웨어 탐지 및 대응

그림 1은 DDoS 탐지용 CNN과 랜섬웨어 탐지용 GRU를 병렬로 결합한 하이브리드 모델을 나타낸다. 송신자는 채널 A와 B를 통해 데이터를 전송하며, 채널 A에서 공격으로 데이터 손상이 발생하면 채널 C로 A와 B의 데이터를 XOR 연산해 전송한다. 수신자는 채널 B의 정상 데이터와 채널 C의 XOR 값을 이용해 채널 A의 손상된 데이터를 복원한다. 제안 기법은 종래 기법 대비 두 공격의 탐지 정확성 향상 및 사용자의 가용성을 개선한다.

3. 성능 평가 및 분석

본 연구에서는 단일 모델 CNN와 GRU을 종래 모델[3]로 선정하여 제안하는 하이브리드 모델의 탐지 정확도를 비교하였고, XOR 기법을 적용한 제안 모

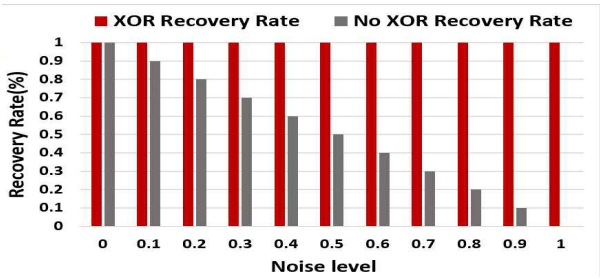
델과 적용하지 않은 종래 모델의 데이터 복원율을 비교하였다. 복원율은 원본 채널 A와 복구된 채널 간의 비트 일치 비율로, 전체 비트 중 동일한 값의 비율로 계산하였다. 실험에서는 DDoS와 랜섬웨어 공격을 탐지하기 위해 머신러닝 데이터인 CICIDS2017과 Ransomware Detection Dataset을 사용하였고, A, B, C 3개 채널로 통신하는 네트워크 환경을 가정했다. 채널 A에서는 DDoS 및 랜섬웨어 공격으로 데이터 손상이 발생하며, 데이터의 비트를 무작위로 선택해 비트 값을 반전시켜 구현하였다. 각 채널에서 송수신하는 데이터 길이는 10비트로 총 1,000개의 샘플 데이터를 생성하고, 공격으로 인한 노이즈 강도는 0부터 100%까지 10% 단위로 증가시켰다. 모든 실험은 100회 반복 시뮬레이션을 수행했다. 그림 2는 Epoch에 따른 CNN, GRU, 및 Hybrid 모델의 검증 정확도 및 DDoS, 랜섬웨어 탐지 시간 CNN, GRU 모델의 학습 시간을 나타낸 그래프이다.



(그림 2) 모델별 검증 정확도 및 DDoS, 렌섬웨어 탐지 학습 시간 비교

그림 2에 따르면, 랜섬웨어 탐지의 경우 GRU가 CNN 대비 18.36% 높은 정확도를 보였다. 랜섬웨어 특성상 점진적인 피해 양상을 보이므로, 실시간 탐지보다는 탐지 정확도가 중요한 요소로 작동하여 GRU가 랜섬웨어 탐지에 적합하다. DDoS 탐지의 경우 CNN이 GRU보다 평균 탐지 정확도는 3.03% 낮지만, 평균 4.207배 짧은 학습 시간이 소요되었다. DDoS는 짧은 시간 내 대량의 트래픽을 유발하므로 빠른 탐지가 중요하므로 정형화된 패턴 인식에 효과적인 CNN이 DDoS 탐지에 적합하다. 따라서 DDoS 탐지에는 CNN, 랜섬웨어 탐지에는 GRU가 효과적이라 판단되어, 이에 본 연구에서는 두 가지 공격 유형에 대한 탐지 성능을 향상시킬 수 있는 CNN, GRU 모델을 결합한 하이브리드 모델을 제안한다.

그림 3은 제안 모델과 종래 모델의 복원율을 나타낸 그래프이다.



(그림 3) XOR 기법 적용 여부에 따른 복원율 비교

제안 모델의 경우 채널 A의 복원율이 100%로 유지됨을 확인할 수 있다. 반면, 종래 모델은 노이즈 강도 비율이 증가함에 따라 복원율이 감소했다.

5. 결론

종래 기법은 단일 유형의 공격 탐지에만 초점을 맞추었고, 공격 대응이 미흡했다. 이러한 한계점을 해결하고자 DDoS 및 랜섬웨어 탐지를 위한 하이브리드 모델과 탐지 이후 복구를 위한 XOR 기반 데이터 복원 기법을 제안했다. 실험 결과에 따르면 제안한 하이브리드 모델은 종래 모델인 CNN, GRU 대비 각각 평균 17.3%, 2.36% 정확도가 향상되었으며, XOR 기법을 적용한 경우, 채널 A에서 복원율이 100%로 사용자의 가용성을 유지할 수 있다.

Acknowledgement

본 논문은 2025년도 과학기술정보통신부 및 한국연구재단의 중견연구 (RS-2025-00518150)와 과학기술정보통신부 및 정보통신기획평가원의 ICT혁신인재4.0 사업의 연구결과로 수행되었음 (No. IITP-2022-RS-2022-00156310)

참고문헌

[1] Cloudflare, "DDoS Threat Report for 2024 Q4," Cloudflare Radar, Jan. 2025. [Online]. Available: https://radar.cloudflare.com/reports/ddos-2024-q4.

[2] S. -J. Lee, H. -Y. Shim, Y. -R. Lee, T. -R. Park and I. -G. Lee, "Ransomware Detection Using Open-source Tools," ICACT, Korea, Republic of, 2022, pp. 1385-1391

[3] S. Sharma, H. Kumar, D. C. Sati, P. Kumar, and M. Monika, "CyberShield: A Hybrid CNN-GRU Model for Intelligent DDoS Attack Recognition," Proceedings of the 2024 INNOCOMP, Greater Noida, India, 2024, pp. 376 - 382.

[4] M. Mittal, K. Kumar, and S. Behal, "Deep learning approaches for detecting DDoS attacks: a systematic review," Soft Computing, vol. 27, pp. 13039 - 13075, 2023.