AI 기반 이상행위 및 위협징후 탐지에 대한 동향

김태훈[1], 김수현[2], 이임영[2]
[1]순천향대학교 소프트웨어융합학과 박사과정
[2]순천향대학교 컴퓨터소프트웨어공학과 교수
20134101@sch.ac.kr, kimsh@sch.ac.kr, imylee@sch.ac.kr

Trends in Anomaly and Threat Detection on AI-based

Taehoon Kim[1], Su-Hyun Kim[2], Im-Yeong Lee[2]
[1]Dept. of Software Convergence, Soonchunhyang University
[2]Dept. of Computer Software Engineering, Soonchunhyang University

요 약

현대 사이버 위협의 복잡성이 증가함에 따라 이상행위 및 위협징후 탐지를 위한 AI 기반 기술 사용이 크게 증가하고 있다. 본 논문은 AI 기반 이상행위 및 위협징후 탐지의 최신 동향을 분석하여 AI 기반 이상행위 탐지 시스템의 탐지 정확도를 높이는 메커니즘과 위협징후 탐지 속도를 가속화하는 기술에 대해 논의한다. 연구 결과에 따르면, AI 기반 모델은 기계 학습 알고리즘의 발전, 보다 다양한 학습 데이터셋, 최적화 기술 덕분에 정확도와 속도 모두에서 개선을 보이고 있다. 본 논문은 현재의 동향과 이들이 AI 기반 보안 시스템의 미래에 미치는 영향을 종합적으로 다루고자 한다.

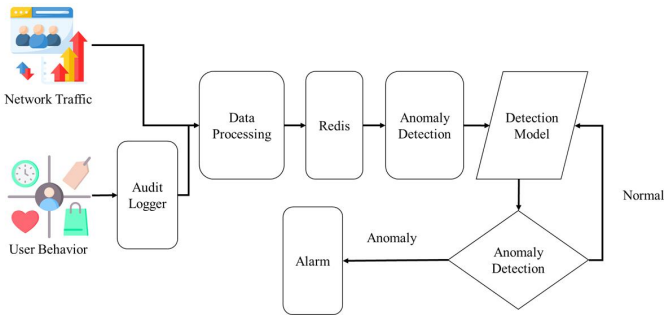
1. 서론

사이버 위협이 점점 더 정교해짐에 따라 전통적인 이상행위 및 위협징후 탐지 방법은 점차 한계를 드러내고 있다. 오늘날의 디지털 환경의 요구를 충족시키기 위해, AI(Artificial Intelligence) 기반 접근법이 강력한 대안으로 떠오르고 있으며, 더 높은 정확도와 더 빠른 탐지 속도의 가능성을 제시하고 있다. 이러한 AI 기반 시스템은 기존의 규칙 기반 시스템이 간과할 수 있는 패턴과 이상 징후를 감지할 수 있다. AI를 보안에 활용하는 사례가 증가함에 따라, 본 논문은 그 중에서도 두 가지 중요한 요소, 즉 이상행위 탐지의 정확성과 위협징후 탐지 속도에 초점을 맞추고 있으며, 이는 위험을 최소화하고 공격에 효과적으로 대응하는 데 중요하다. 최근 연구를 분석함으로써 AI 기반 보안 솔루션의 최신 동향을 탐구하고, 현재 상태에 대한 통찰을 제공한다.

2. 배경

2.1 AI 기반 이상행위 탐지

AI 기반 이상행위 탐지는 그림 1과 같이 시스템, 네트워크, 또는 사용자 활동에서 일반적인 패턴을 벗어난 비정상적인 행위(이상)를 자동으로 감지하는 기술이다[1]. 이러한 이상은 종종 새로운 사이버 공격, 네트워크 침입, 데이터 유출 시도를 시도할 수 있으며, 전통적인 규칙 기반 시스템은 감지하기 어려운 경우가 많다.



(그림 1) AI 기반 이상행위 탐지 워크플로우

AI 기반 이상행위 탐지는 기계 학습 알고리즘을 사용하여 대규모의 정상 데이터를 학습하고, 이를 바탕으로 이상치(Outliers)나 예측 불가능한 변동을 탐지한다. 이를 통해 알려지지 않은 공격, 제로데이 공격, 또는 잠재적 위험을 빠르게 식별할 수 있다.

그러나 AI 기반 이상행위 탐지의 정확도는 학습 데이터의 품질, 알고리즘 성능, 데이터 다양성 등 여러 요인에 따라 달라진다. 정확도를 높이는 연구는 오탐지(False Positives)를 줄이고, 정교한 공격을 탐지하는 데 중요한 역할을 한다.

2.2 AI 기반 위협징후 탐지

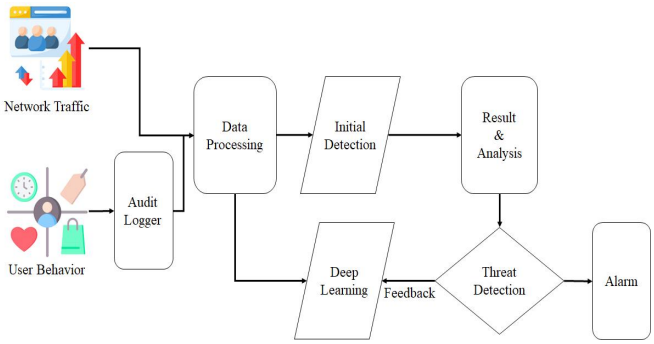
AI 기반 위협징후 탐지는 그림 2와 같이 네트워

<표 2> AI 기반 이상행위 탐지 관련 연구 비교 분석표

	훈련 모델	최대 탐지 정확도	최대 탐지 시간
[2]	LSTM, GRU, CNN, MLP	99.87%(LSTM 모델)	LSTM: 약 4시간, GRU: 약 4시간, CNN: 약 4시간, MLP: 약 6.5시간
[3]	DNN, SAE, RBM, SDAE	94.6%(DNN 모델)	-
[4]	Stacked LSTM, Bidirectional LSTM	99.15% (mean_squared_error 기반 손실함수의 Stacked LSTM)	-

<표 3> AI 기반 위협징후 탐지 관련 연구 비교 분석표

	훈련 모델	최대 탐지 정확도	최대 탐지 시간
[5]	Ensemble FS, DNN, CNN	87.2%(DNN 모델)	Ensemble FS: 159.632초, DNN: 142.672초
[6]	OML	98%(OML)	OML: 2초 이내



(그림 2) AI 기반 위협징후 탐지

크나 시스템에서 발생하는 공격의 전조 또는 징후를 실시간으로 감지하고 이를 위협으로 분류하는 기술이다. 위협징후란 악성 소프트웨어, 의심스러운 활동, 허가되지 않은 접근 시도 등 보안 침해가 발생할 가능성을 예고하는 경고 신호를 말한다. AI 기반 탐지 시스템은 이 징후들을 학습하고 탐지하는 역할을 수행한다.

AI 기반 위협징후 탐지는 사전에 정의된 악성 패턴(서명 기반 탐지)뿐만 아니라, 기계 학습 알고리즘을 활용한 비정형 데이터 탐지(행동 기반 탐지)를 통해서도 이루어진다. 이를 통해 알려진 공격뿐만 아니라 새롭게 등장하는 공격도 감지할 수 있다.

특히, 위협징후 탐지 속도는 신속한 대응에 필수적이며, 탐지 속도가 느리면 심각한 피해로 이어질 수 있다. 따라서 AI 기반 시스템의 탐지 속도를 최적화하는 연구는 보안 사고를 최소화하고 정교한 대응을 가능하게 한다.

3. 관련 연구 분석

3.1 AI 기반 이상행위 탐지 정확도

Al-Shabi et al. [2]은 무단 접근 및 의도치 않은 수정 문제와 같은 IoT(Internet of Things) 보안 위협 요소들을 제시하였다. 이를 해결하기 위해 LSTM(Long Short-Term Memory) RNN(Recurrent Neural Networks)을 채택하여 IoT 센서 데이터셋을 기반으로 높은 정확도의 비정상 행동 탐지 모델을 구축하였다.

Abusitta et al. [3]은 IoT 데이터 무결성 및 서비스 가용성을 위협하는 문제(IoT 이질성, 노이즈 등)를 탐지하기 위해 노이즈 제거 오토인코더를 활용한 딥러닝 기반 탐지 프레임워크를 제안하였다.

Girish et al. [4]은 클라우드에서 전통적인 모니터링 방법에 대한 이상행위 탐지에 대한 한계를 제시하였다. 이를 해결하기 위해 순차적 데이터 분석에 효과적인 LSTM 네트워크를 사용하여 자동화된 이상 탐지 방법을 제안했다.

3.2 AI 기반 위협징후 탐지 속도

Saha et al. [5]은 DDoS 공격으로부터 안전하게 사이버 보안을 제공하기 위해 DDoS 공격을 식별하고자 하였다. DDoS 공격을 분류하기 위해 최적의 피처 세트를 선택해야 했으며, 이는 딥러닝 기반 15개의 개별 피처 세트, 앙상블 피처 세트, 원래 피처 세트에 대한 포괄적인 비교 분석을 수행하여 다수결 투표 기법 기반 최적의 피처 세트를 식별했다.

Akbar et al. [6]은 APT(Advanced Persistent Threats) 공격을 탐지하기 위해 호스트 시스템에서

이벤트 데이터를 수집하기 위해 출처 그래프를 활용하고, OML(Online Metric Learning) 기반 딥러닝을 이용하여 위협징후를 탐지한다.

4. 결론

AI 기반 이상행위 및 위협징후 탐지 기술은 점점 더 정교해지는 사이버 공격에 효과적으로 대응하기 위한 핵심적인 도구로 자리 잡고 있다. 본 논문은 이상행위 탐지의 정확성과 위협징후 탐지 속도를 중심으로 AI 기반 보안 솔루션의 최신 동향을 분석하였다. 연구 결과, 기계 학습 알고리즘의 발전과 다양한 학습 데이터셋의 활용, 최적화 기술이 AI 시스템의 성능을 크게 향상시키는 것으로 나타났다. 특히, 이상행위 탐지 정확도는 오탐지를 줄이고, 더욱 정교한 공격을 식별하는 데 중요한 역할을 하며, 위협징후 탐지 속도는 신속한 대응을 가능하게 하여 피해를 최소화한다.

향후 AI 기반 보안 시스템의 발전은 이러한 성능 향상 요소에 대한 지속적인 연구와 더불어, 시스템의 해석 가능성 및 적응성을 높이는 방향으로 이루어져야 할 것이다. 또한, 새로운 사이버 위협에 대응할 수 있는 능력과 실시간 분석을 위한 더 빠르고 정확한 탐지 기술이 요구된다. 결론적으로, AI 기반 보안 솔루션은 점차 복잡해지는 위협 환경에서 중요한 역할을 수행할 것이며, 이를 효과적으로 활용하기 위한 연구와 개선이 필수적이다.

Acknowledgment

본 연구는 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원(RS-2022-00167197, 스마트시티 구축을 위한 지능형 5G/6G 핵심 인프라 기술 개발)과 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 3단계 산학연협력 선도대학 육성사업(LINC 3.0)의 지원을 받아 수행된 연구임(과제번호 : 1345356224).

참고문헌

[1] Zhao, Y., et al. "Network Anomaly Detection by Using a Time-Decay Closed Frequent Pattern," Information, Vol. 10, No. 8, 262, 2019.

[2] Al-Shabi et al. "Using deep learning to detecting abnormal behavior in internet of things", International Journal of Electrical and Computer Engineering, Vol. 12, No. 2, pp. 2108-2120, 2022.

[3] Abusitta, Adel, et al. "Deep learning-enabled anomaly detection for IoT systems", Internet of Things, Vol. 21, 100656, 2023.

[4] Girish, L., and Rao, S.K. "Anomaly detection in cloud environment using artificial intelligence techniques", Computing, Vol. 105, No. 3, pp. 675-688, 2023.

[5] Saha, Sajal, et al. "Towards an optimal feature selection method for AI-based DDoS detection system.", 2022 IEEE 19th Annual Consumer Communications & Networking Conference, Virtual Conference, 2022, pp. 425-428.

[6] Akbar, K. Ashrafi, et al. "Advanced Persistent Threat Detection Using Data Provenance and Metric Learning.", IEEE Transactions on Dependable and Secure Computing, Vol. 20, No. 5, pp 3957-3969, 2022.