AI 기반 보안관제의 문제점 고찰

안중현* · 최영렬 · 백남균

부산외국어대학교

A Study on the Problems of AI-based Security Control

Jung-Hyun Ahn* · Young-Ryul Choi · Nam-Kyun Baik

Busan University of Foreign Studies

E-mail : endrk723@naver.com / dudfuf261@naver.com / namkyun@bufs.ac.kr

요 약

현재 보안관제 시장은 AI기술을 기반으로 하여 운영 중이다. AI를 사용하는 이유는 보안장비간 대량으로 발생하는 로그와 빅데이터에 대해 이를 탐지하기 위해 사용하고, 시간적인 문제와 인적인 문제를 완화하기 위해서 이다. 하지만 AI를 적용함에도 문제는 여전히 발생하고 있는 중이다. 보안관제 시장은 이 논문에서 소개하는 문제점 말고도 많은 문제점과 대응하고 있으며, 본 논문은 다섯 가지의 문제점을 다루고자 한다. ‘AI 모델 선정’, ‘AI 표준화 문제’, ‘빅데이터의 정확성 및 신뢰성’, ‘책임소재의 문제’, ‘AI의 타당성 부족’ 등 보안관제 환경에 AI기술을 적용함에도 발생하는 문제점을 고찰하고자 한다.

ABSTRACT

Currently, the security control market is operating based on AI technology. The reason for using AI is to detect large amounts of logs and big data between security equipment, and to alleviate time and human problems. However, problems are still occurring in the application of AI. The security control market is responding to many problems other than the problems introduced in this paper, and this paper attempts to deal with five problems. We would like to consider problems that arise in applying AI technology to security control environments such as ‘AI model selection’, ‘AI standardization problem’, ‘Big data accuracy’, ‘Security Control Big Data Accuracy and AI Reliability’, ‘responsibility material problem’, and ‘lack of AI validity.’

키워드

Security control, AI, Big data , Detection , Problem

Ⅰ. 서 론

현재 보안관제 시장은 AI기술을 기반으로 하여 운영 중이다. 그 이유는 보안장비간 서로 발생하는 로그들이 있고, 이러한 로그 데이터들은 하루에도 백만 이상으로 발생하고, 데이터에 대해 유해한 정보가 있는지 탐지 및 분석을 거쳐야 한다. 하지만 인원의 수는 제한적이고, 보안관제 업무상 24시간 동안 모니터링을 통하여 데이터를 감시해야하고, 실시간으로 수집한 데이터를 빠른 시간 안에 분석해야하기 때문에 시간적인 제약도 따라온다. 때문에 보안관제에 AI기술을 도입하면서 사람이 해결하기 어려운 문제를 처리하여 완화 되었지만, 이러한 기술을 사용에도 불구하고 문제점이 있는 상황이다. 본 논문에서는 기존의 보안관제 환경에서 AI기술을 적용하면서 발생하는 문제들을 나열하고, 고찰하고자 한다.

Ⅱ. 관련현황

국내에서 보안관제의 운영 방식은 보안 장비들간 발생하는 보안로그들을 보안정책을 이용하여 데이터를 탐지 및 분석하는 방식이다. 보안로그란

* speaker

방화벽, DDoS 장비, IDS(침입탐지시스템), IPS(침입차단시스템), WAF(웹방화벽) 등의 장비에서 탐지 된 네트워크 패킷의 일부이다. 현재 보안관제 시장에서 운영되고 있는 AI기반 기술들은 데이터 수집, 데이터 전처리, 학습 및 탐지 등 외에도 다양한 기술들이 쓰이고 있다. 데이터 수집은 위와 같이 대량의 데이터를 기관이 정보를 수집하여 미래에 같은 공격 데이터가 침입할 경우 그 공격에 대한 방안을 마련하기 위해 수집한다. 데이터 전처리는 AI가 스스로 학습하는 과정에 있어 학습에 대한 데이터의 품질 및 성능을 높이기 위하여 수집된 샘플링, 변환 등을 위해 유형을 정의하고, 유사 분석 등 다양한 단계를 적용한다. 수집된 데이터에서 결과값을 예측할 수 있는 특징을 추출하는데 이때 특징을 피처라고 부른다. 앞의 과정을 피처 추출 이라고 부른다. 학습 및 탐지는 정、오탐의 식별과 이상행위 탐지 등 목표로 하는 모델을 생성하는 단계이다. 과정은 앞의 전처리된 데이터를 입력받고 학습 모듈, 탐지 모듈을 사용해 선택된 머닝러신 및 딥러닝 등 알고리즘을 접하여 새로운 모델을 생성하는 것이다. 이렇듯 보안관제 환경에 AI기술을 기반으로 하여 많은 제품들이 개발되고 있지만, 실제로 업계에서 사용되고 있지 않고 있다. 많은 문제점들 중 다섯 가지의 문제점을 소개하고자 한다. ‘AI모델 선정’, ‘AI의 표준화의 문제’, ‘빅데이터의 정확성 및 신뢰성’, ‘책임소재의 문제’, ‘AI의 설명 부족’ 등 여러 문제점들을 살펴보겠다. [1][2]

III. 보안관제에 AI기술 적용시 문제점

위와 같이 보안관제 환경에 AI기술을 적용할 시, 발생하는 다섯 가지의 문제점을 살펴보겠다.

1. AI 모델 선정

보안관제 환경에서 AI기술 적용은 시간적 제약과 인적부담을 덜어주기 위해 사용되고 있다. 하지만 AI의 모델을 선정하는데 시간적으로 많은 소요가 걸린다는 문제이다. 이유는 보안관제 기기마다 사용하는 알고리즘이 다르기 때문에 가장 적합한 모델을 찾기에는 모델이 제한적이며, 기기를 호환할 수 있는 모델이여야 한다. 그리고 모델의 역할은 분석대상인 데이터에 대해 이상행위 탐지를 하는데 정상인 데이터와 사이버 위협 데이터를 비교하여 정상 수치에서 오차가 클 경우 공격 데이터라 판단한다. 좋은 모델일수록 정상적인 데이터를 보다 많이 학습하여 현 보안관제에 많은 도움을 줄 수 있지만, 보안관제에 사용하는 장비들을 모두 호환하며 가장 좋은 성능치의 모델을 선정하기 위해 다른 제품들과 비교하는 과정을 거치며 많은 시간이 소요된다.

2. AI 표준화 문제

보안관제에 AI기술을 적용하기 전, 국내보안관제 시장은 각 업체 및 기업마다 관리 및 운영방식은 달랐다. 이유는 각 기관의 역량에 따라 보안기술이 달라지기 때문이다. 기관의 역량이 높을수록 새로운 사이버 위협 데이터에 대한 대응기술이 보다 발전되어 있을 것이다. 고객의 입장에서 대응기술의 수준이 다른 기관보다 높은 것을 선호하기에 한 쪽 기관을 더 선택하게 될 것이고, 이것은 곧 경제적인 문제가 발생한다. 이처럼 기관간 경쟁구도가 이루어지기 때문에 사이버 위협 정보에 대한 공유와 소통은 잘 이뤄지지 않으며, 현재 AI를 적용한 보안관제 환경에서도 유지되고 있어 문제점을 유발하는 상황이다.

3. 보안관제 빅데이터의 정확성 및 AI 신뢰성

빅데이터를 통한 학습하는 과정에서 일어나는 문제이다. 정확하고 올바른 데이터들을 바탕으로 학습을 진행시 데이터에 대한 정탐률(True Positive)은 오르지만 거짓정보를 옳다고 판단하여 학습할 경우, 미탐률(False Negative)이 증가하여 문제가 발생한다. 관련된 사례로 지난 2016년 마이크로소프트(MS)사에서 개발한 AI로봇 테이는 사람들과 소통하기 위해 만들어진 AI 채팅봇이다. 테이는 구글의 바둑 프로그램 알파고의 기술인 신경망이라고 불리는 AI기술을 가반으로 제작됐다. 사람들과 대화를 하면서 단어 사용법, 질문방식, 대답의 방법 등 패턴을 익히고 스스로 학습한다. 학습하는 과정에서 어느 익명게시판을 통하여 테이가 인종차별 발언을 하도록 훈련시키자는 글이 올라왔고, 이러한 집단에 의해 잘못된 데이터를 학습하게 되었다. 후에는 다른 일반사람들과 대화 시, 욕설이나 차별적인 발언 등 물의를 일으켜 MS는 문제가 된 테이를 운영을 중지하고 삭제한 사례가 있다. 이처럼 AI가 올바른 학습을 위해서 학습하는 데이터는 데이터 자체가 순수하고, 정제되어 있어야 AI가 학습을 통하여 질문에 대한 해답을 할 때, 올바른 판단을 내릴 수 있다. 하지만 보안에 AI기술을 적용함에 있어 지나친 신뢰는 오히려 독이 될 수 있다. 위의 빅데이터 학습에 있어 거짓정보나 편향된 데이터를 학습하여 그것을 옳은 판단이라고 판단했을 경우 문제가 발생한다. 위의 테이로봇에 대해 학습하는 과정중, 사람의 개개인의 생각이 다르기 때문에 어떠한 이벤트에 대해 판단이 달라질 수 있다. 다시말해 어떤 업체에 보안사고가 일어 났고, 방안이 여러개가 존재한다고 가정해보자. 각 방안마다의 장점과 단점이 존재하고, 어느 것이 해당 이벤트에 대해 적합한지 사람마다 생각은 다를 것이다. 이렇듯 한 이벤트에 대해 여러 판단이 나오고 AI는 편향된 데이터를 학습하게 되어 이벤트에 대해 명확한 판단을 내리지 못하는 문제

가 발생된다. 그리고 보안장비간 발생하는 대량에 데이터에 대해 사이버 위협을 초례할 데이터임을 확인하는 과정도 많은 시간적 요소가 되며, 이러한 AI를 실무에 적용하기에 정확도가 불안전하여 AI에 대한 신뢰성이 부족한 상황이다. [3]

4. 책임소재의 문제

AI의 책임소재의 문제 또한 보안관제 환경에 일어날 수 있는 문제이다. AI는 빅데이터와 실시간으로 수집되는 정보를 바탕으로 학습을 하고 서버에 사이버 위협으로 의심되는 데이터 유입 시, 학습을 토대로 위협 데이터가 맞는지, 아닌지를 판단할 것이다. 만약 AI가 잘못된 판단을 내려 사이버 위협이 서버로 전송되고 보안사고로 이어질 경우, 기관에 피해가 갈 경우 책임은 누구에게로 돌아가는지에 대해서 문제가 발생한다. AI 기술을 발명한 사람의 책임으로 넘어갈 것인지, 보안장비 기기를 제조한 회사의 책임으로 넘어 갈 것인지, 사이버 공격을 받은 회사의 책임으로 넘어갈 것인지 등으로 어느 한 기관의 책임이라고 판단하기에 현재로서는 AI에 관한 법률이 정확히 나오지 않은 상황이라 문제로 다뤄지고 있다. [4]

5. AI의 설명 부족

마지막으로 AI의 설명 부족에 대한 문제점이다.

AI는 확인이 되지 않은 데이터의 유입시, 들어온 데이터에 대하여 사이버 공격 데이터인지, 아닌지에 대해 판단을 할 것이다. AI가 판단하는 기준은 학습된 데이터를 바탕으로 과거에 비슷한 데이터를 학습하고 결과로 내었던 값을 도출한다. 하지만 이 과정에서 우리에게 보여주는 값은 오직 결과값만 보여준다. 즉, AI가 어떠한 이벤트에 대해 어떻게 판단을 하고, 그 과정이 어떻게 나왔는지에 대해 알 수 없는 상황이다. 즉, AI는 해당 이벤트를 받으면 어떠한 알고리즘을 사용하며, 해답을 내놓는 과정이다. 하지만 그 과정은 현재 사람들에게 보여주지 않고 이벤트에 대해 AI가 판단하여 내놓은 결과값만을 보여주는 것이다. 설명 가능한 AI를 사용함으로써 기관은 이해하는 당사자가 AI모델의 작동과 어떠한 알고리즘을 적용했는지 이해할 수 있도록 도와줄 수 있지만, 현재는 그렇지 못하여 아쉬운 상황이다. [5]

Ⅳ. 결 론

보안관제에 AI기술을 적용하면서 시간적, 인적 제한을 완화시켜주지만, AI기술을 적용함에 발생하는 문제점들을 살펴보았다. 각 문제점들을 보완하고 해결방안을 적용시킨다면, 업무의 효율성은 향상될 것이고, 다가올 새로운 사이버 위협으로부터 조치 및 예방이 가능한 환경이 마련될 것을 전망한다.

Acknowledgement

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 ICT혁신인재4.0사업의 연구결과로 수행되었음(IITP-2021-2020-0-01825)

References

[1] Y. G. Ju, S. H. Kim, C. J. Woo, D. J. Ryu, S. H. Kang, “A Practical case Study for Maximizing the True Positive rate in Implementing a Security Monitoring Systems Based on A.I”, *Journal of Collection of papers from the Korean Society of Information and Communications*, Korea, pp. 1165-1166, Jan. 2019

[2] I. O. Jung, C. S. Cho, J. W. Ji, “Problems with the cybersecurity control system and Current status of machine learning application technology”, *Journal of Information Protection*, Vol.31, No.3, Korea, pp. 15-17, Jun. 2021

[3] Seoul newspaper website [Internet]. Available : https://www.seoul.co.kr/news/newsView.php?id=20160325800054

[4] Igloosecurity website [Internet]. Available : http://www.igloosec.co.kr/BLOG_%EC%9D%B8%EA%B3%B5%EC%A7%80%EB%8A%A5%EC%9D%98%20%EC%9E%A5%EC%A0%90%EA%B3%BC%20%ED%95%9C%EA%B3%84?searchItem=&searchWord=&bbsCateId=0&gotoPage=8

[5] IBM website [Internet]. Available : https://www.ibm.com/kr-ko/watson/explainable-ai