Autoencoder를 사용한 SWaT 데이터 상의 이상 탐지에 관한 연구

권희용, 이문규*

*인하대학교

heeyong.kr@gmail.com, *mklee@inha.ac.kr

A Study on the Anomaly Detection on SWaT dataset using Autoencoder

Hee-Yong Kwon, Mun-Kyu Lee*

*Inha Univ.

요 약

본 논문은 현대화된 수처리 시스템에서 발생할 수 있는 사이버공격에 의한 이상 동작 탐지 기법의 성능 평가를 위해 만들어진 공개 데이터셋 중 하나인 SWaT 데이터셋을 대상으로 비지도 학습 방법 중 하나인 Autoencoder의 다양한 모델을 설계 및 적용함으로써 이상 동작을 탐지하고 모델의 성능을 측정하였다.

Ⅰ. 서 론

최근의 산업 제어 시스템(ICS)은 시스템을 관리하는 다양한 기기들이 네트워크에 연결되고 있다. 이에 따라 제어 시스템은 효율적으로 관리 및 제어되나, 네트워크 연결에 따른 다양한 공격 경로가 생겨나게 되었다. 이에 사이버공격으로 인한 ICS의 오동작 또는 이상 동작을 탐지하는 연구의 필요성이 증가하였고, 많은 연구진들에 의해 이상 동작 탐지에 관한 연구들이 진행되었다. Secure Water Treatment(SWaT) 데이터셋[1]은 수처리 시설의 공개 데이터셋으로서, 공개 데이터셋이 부족한 ICS 보안 분야의 다양한 연구에 사용되고 있다[2, 3]. 본 논문에서는 ICS 보안을 위한 하나의 사례 연구로 SWaT 데이터셋을 대상으로 이상 탐지 연구를 진행하였다. 이상 동작의 탐지를 위해 세 가지 종류의 Autoencoder 모델을 사용하였으며, 각 모델의 이상 탐지 성능을 비교한다.

Ⅱ. 본론

SWaT 데이터셋은 Singapore University of Technology and Design의 기관인 iTrust에서 공개한 데이터셋으로 7일간의 정상 동작에서 수집된 데이터와 이상 동작을 포함하는 4일간의 데이터로 구성되며, 테스트베드에서 생성되는 센서 및 actuator의 데이터와 송·수신되는 네트워크 트래픽 데이터를 포함한다. 본 연구에서는 SWaT 데이터셋의 센서 및 actuator 데이터를 이상 동작 탐지의 입력으로 사용하였으며, 딥러닝 알고리즘은 Intel Xeon Gold 6242 CPU @ 2.8GHz, 256GB RAM, 그리고 NVIDIA TITAN RTX GDDR6 24GB GPU를 탑재한 기기 상에 Python에서 가장 많이 사용되는 딥러닝 라이브러리 중 하나인 Keras를 활용한 Autoencoder를 설계 및 사용하였다. 실험에 사용된 Autoencoder는 총 세 종류로, 입력 데이터를 재구성(Reconstruction)하는 모델, 입력 데이터를 바탕으로 다음 데이터를 예측(Prediction)하는 모델, 그리고 재구성과 예측을 동시에 사용하는 합성(Composite) 모델이 사용되었다. 또한 Autoencoder는 입력 및 출력 레이어를 제외한 5개의 히든 레이어를 포함하며 각 레이어는 64, 32, 16, 32, 64 개의 노드를 가지는 Dense 레이어로 구성되었고, 활성 함수로 Hyperbolic tangent, 손실 함수로 Mean squared error, 그리고 모델 최적화를 위한 optimizer로써 Adam optimizer[4]가 사용되었다. 설계된 Autoencoder를 사용한 이상 탐지 성능을 평가하는 지표로써 Precision, Recall, F1-score를 사용하였고, 각 모델을 사용하여 SWaT 데이터셋의 이상을 탐지한 결과는 아래 표와 같다.

	Precision	Recall	F1-score
Reconstruction	**0.985**	0.596	0.742
Prediction	0.964	0.651	**0.777**
Composite	0.731	**0.804**	0.766

Ⅲ. 결론

본 논문에서는 산업 제어 시스템을 대상으로 하는 보안 위협에 대응하기 위한 사례 연구의 하나로써, SWaT 데이터 상의 이상을 탐지하였다. 이상 탐지를 위해 잘 알려진 비지도 학습 방법 중 하나인 Autoencoder가 사용되었으며, Precision의 측면에서는 재구성 모델이, Recall 측면에서는 합성 모델이, 그리고 F1-score 측면에서는 Prediction 모델이 가장 좋은 성능을 보임을 확인할 수 있었다.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 글로벌핵심인재양성지원사업의 연구결과로 수행되었음 (2020-0-01540)

참 고 문 헌

[1] Goh, Jonathan, et al. "A dataset to support research in the design of secure water treatment systems." International conference on critical information infrastructures security. Springer, Cham, 2016.

[2] Wang, Chao, et al. "Anomaly Detection for Industrial Control System Based on Autoencoder Neural Network." Wireless Communications and Mobile Computing 2020 (2020).

[3] Mieden, Philipp, and Rutger Beltman. "Network Anomaly Detection in Modbus TCP Industrial Control Systems." (2020).

[4] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014).