JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 24, No. 8, pp. 1929-1936, Aug. 2023

Check for updates

비지도 학습 기반 오토인코더를 사용한 내부자 이상 행위 탐지 방법

심 기 천[1] · 김 강 석[2*]
[1]아주대학교 지식정보공학과 석사과정
[2*]아주대학교 사이버보안학과 교수

Insider Anomaly Behavior Detection Method Using an Unsupervised Learning-Based Autoencoder

Ki-Chun Sim[1] · Kangseok Kim[2*]

[1]Master's Course, Department of Knowledge Information Engineering, Ajou University, Suwon 16499, Korea
[2*]Professor, Department of Cyber Security, Ajou University, Suwon 16499, Korea

[요 약]

본 논문에서는 내부자의 행위를 기록한 시계열 기반 로그 데이터를 사용하여 내부자 이상 행위 탐지 방법을 연구한다. LSTM 기반 잡음 제거 오토인코더(LSTM-DAE) 모델을 개발하고, 이를 활용하여 유용한 시퀀스 정보를 담고 있는 잠재 벡터를 추출하였다. 그리고 추출한 잠재 벡터를 이상 탐지 알고리즘인 LOF와 IF에 입력하여 내부자 이상 행위 탐지 방법의 성능을 평가하였다. 여러 가지 성능 평가 지표를 사용하여 모델의 실효성을 검증한 결과, 5차원인 잠재 벡터를 사용하면 시퀀스 길이가 짧을수록 재현율이 높게 나온 것을 확인할 수 있었고, 7차원인 잠재 벡터를 사용하면 시퀀스 길이에 상관없이 재현율이 높게 나온 것을 확인할 수 있었다. 또한 비정상 행위 샘플 수를 일정하게 유지하면서 정상 행위 샘플 수가 증가할수록 정밀도가 하락하는 것을 확인할 수 있었다.

[Abstract]

Herein, we study an insider anomaly behavior detection method using time-series-based log data that records insider behavior. We developed a long-short-term-memory-based denoised autoencoder model and extracted latent vectors containing useful sequence information from the autoencoder. The performance of the insider anomaly detection method was further evaluated by inputting the extracted latent vectors to anomaly detection algorithms—Local Outlier Factor and Isolation Forest. By verifying the effectiveness of the model using various performance evaluation indicators, via the coding vector (dimension: 5), it was confirmed that the shorter the sequence length, the higher the recall, and using the coding vector (dimension: 7), the higher the recall regardless of the sequence length. Furthermore, while keeping the number of abnormal behavior samples constant, it was confirmed that the precision decreased as the number of normal behavior samples increased.

색인어 : 내부자 이상 행위 탐지, 인공지능, 머신러닝, 딥러닝, 오토인코더

Keyword : Insider Anomaly Behavior Detection, Artificial Intelligence, Machine Learning, Deep Learning, Autoencoder

http://dx.doi.org/10.9728/dcs.2023.24.8.1929

Received 21 July 2023; **Revised** 03 August 2023
Accepted 03 August 2023

***Corresponding Author; Kangseok Kim**

Tel: +82-31-219-2496
E-mail: kangskim@ajou.ac.kr

Ⅰ. 서 론

정보보안을 위해 이상 행위에 대응하는 것이 중요해지면서 이상(Anomaly)을 식별하고 분석하는 이상 탐지 연구에 대한 필요성이 증가하였다. 초기에는 사전에 정의된 이상 패턴이나 시그니처를 이용한 연구가 주를 이루었지만, 새로운 이상 패턴이나 변형된 공격에 대응이 어려워 머신러닝(Machine Learning) 기반 이상 탐지 연구가 주목받게 되었다.

일반적으로 지도 학습(Supervised Learning) 기반 이상 탐지 방법을 사용할 경우 성능은 좋은 편이지만, 학습을 위해 레이블 된 충분히 많은 양의 데이터가 필요하고, 정상 데이터와 비정상 데이터를 분류하는 작업에 많은 자원이 소모된다. 또한 매우 적은 양의 비정상 데이터로 인해 다수를 차지하는 클래스에 편향된 학습이 이루어져 소수의 비정상 클래스에 대한 분류 성능이 저해되기 때문에 실제 환경에서 사용하기 어렵다[1]. 더욱이 이상 행위는 정교해지고 있어 탐지하기가 더 어려워지고 있다[2]. 기존의 탐지 시스템은 이러한 이상 행위의 고도화된 특성으로 인해 새로운 유형의 이상 행위를 탐지하고 대응하는 데 어려움이 있다. 따라서 최근에는 정상 데이터만을 학습하여 데이터의 내재된 패턴을 바탕으로 이상 탐지를 수행하는 비지도 학습(Unsupervised Learning) 기반 딥러닝(Deep Learning) 알고리즘이 활발히 연구되고 있다.

그러므로 본 연구의 목적은 시계열 기반 내부자의 행위를 기록한 로그 데이터로부터 유용한 시퀀스(Sequence) 정보를 보존하고 있는 잠재 벡터(Latent Vector)를 추출하고, 이를 활용하여 내부자 이상 행위 탐지 모델을 개발하는 데 있다. 본 연구에 사용된 데이터는 딥러닝 기반 내부자 위협 탐지 연구에 많이 활용되는 CMU CERT dataset 이다[3]. 가변 길이로 구성된 데이터를 고정 길이의 임베딩 벡터로 변환한 후, 시퀀스 데이터 학습에 적합한 순환 신경망 알고리즘인 LSTM(Long Short-Term Memory)을 활용한 잡음 제거 오토인코더(Denoising Autoencoder, DAE)에 입력하여 잠재 벡터를 추출하였다. 그리고 추출한 잠재 벡터를 이상 탐지 알고리즘에 입력하여 성능 평가를 수행하였다.

본 논문의 구성은 다음과 같다. 2장에서는 비지도 학습 기반 이상 탐지 연구들을 기술하고, 3장에서는 본 연구에 사용된 데이터를 소개하며, 잡음 제거 오토인코더를 활용한 내부자 이상 행위 탐지 방법을 서술한다. 4장에서는 제안 방법에 대한 실험 결과를 분석하고, 5장에서는 결론 및 향후 연구 방향에 대해 제시한다.

Ⅱ. 관련 연구

최근 비지도 학습을 기반으로 하는 이상 탐지 방법들이 활발히 연구되고 있다. 비지도 학습 기반 이상 탐지 알고리즘인 LOF(Local Outlier Factor)는 데이터의 상대적 밀도를 고려하여 이상 데이터를 판별하고[4], IF(Isolation Forest)는 결정 트리를 사용하여 이상치를 찾으며[5], OC-SVM (One-Class Support Vector Machine)은 초평면을 사용하여 데이터를 구분하는 방식으로 이상 탐지를 수행한다[6]. 이뿐만 아니라 복잡한 데이터를 처리하는 데 효과적인 오토인코더(Autoencoder)[7], Deep SVDD(Deep Support Vector Data Description)[8] 등과 같은 비지도 학습이 적용된 딥러닝 기반 이상 탐지 알고리즘들도 개발되었다. 오토인코더는 인코더(Encoder)와 디코더(Decoder)가 결합한 딥러닝 기반 비지도 학습 방법으로 인코더에서는 입력 데이터의 차원을 낮추어 잠재 벡터를 추출하고, 디코더에서는 추출된 잠재 벡터를 사용하여 입력 데이터와 유사한 데이터를 재구성한다. 또한 잠재 표현(Latent Representation)의 크기를 제한하거나 입력에 잡음을 추가한 후, 원본 입력을 복원시켜 데이터를 효율적으로 표현한다[9].

일반적으로 이상 행위 탐지는 전체 데이터에서 정상으로 분류된 데이터와 다른 형태를 지닌 데이터 혹은 정해진 범위를 벗어나는 의심스러운 패턴을 보이는 데이터를 감지하는 것으로, 딥러닝 기반 이상 행위 탐지 방법이 활발히 연구되고 있다[10]. 입력 데이터와 복원 데이터의 재구성 오차(reconstruction error)에 임계값을 두어 이를 넘어서면 해당 샘플을 이상으로 간주하는 방법이 사용되기도 하지만, 임계값 결정에 사람의 주관적인 판단이 포함되는 단점이 존재하여 이를 극복하기 위해 인코더로부터 생성된 잠재 벡터를 활용하는 이상 탐지 방법이 제안되었다[11]. 또한 Kim 등은 시퀀스 데이터를 처리할 수 있는 순환 신경망 알고리즘인 LSTM을 활용한 오토인코더 모델로 데이터 유출 징후를 탐지하고, 내부자의 직급과 5가지 성격 특성 요소에 따른 페널티를 부가하여 실효성이 있는지 검증하였다[12]. Pantelidis 등은 오토인코더 모델과 2차원으로 축소된 잠재 공간(Latent Space)을 갖는 변이형 오토인코더(Variational Autoencoder) 모델로 내부자 행위를 정상 혹은 비정상으로 분류하고, 모델의 성능을 비교 및 분석하는 연구를 진행하였다[13]. Koutsouvelis 등은 불법 행위를 색상으로 구분하는 등 데이터를 이미지로 변환하여 시각화하고, CNN (Convolutional Neural Network) 알고리즘에 입력하여 도출된 결과로 잠재적 위협을 식별하였다[14].

Ⅲ. 연구 방법

본 연구에서는 시퀀스 데이터를 벡터로 변환하는 임베딩 기법과 비지도 학습 기반 오토인코더 등을 활용하여 내부자 이상 행위 탐지 방법을 개발하였으며, 개발 과정은 크게 세 단계로 나뉜다. 첫 번째 단계에서는 시계열 기반의 데이터를 전처리하고, 원-핫 인코딩(One-Hot Encoding)을 하여 임베딩 벡터로 변환하는 작업을 수행한다. 두 번째 단계에서는 변

환한 임베딩 벡터를 비지도 학습 기반 오토인코더에 입력하여 학습시킨 후, 잠재 벡터를 추출한다. 세 번째 단계에서는 추출한 잠재 벡터를 비지도 기계 학습 기반 이상 탐지 알고리즘인 LOF와 IF에 입력하여 내부자 이상 행위 탐지 모델의 성능을 평가한다. 그림 1은 본 연구에서 제안하는 비지도 학습 기반 내부자 이상 행위 탐지 방법의 전체적인 작업 흐름을 나타낸다.

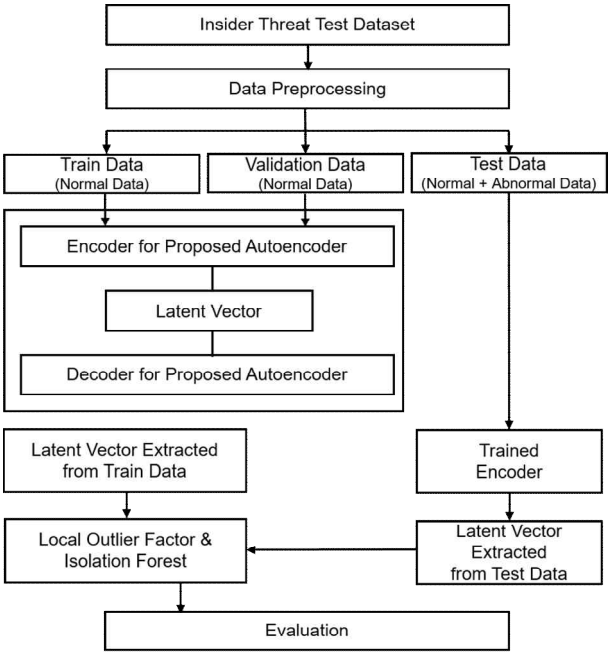

그림 1. 본 연구에서 제안하는 내부자 이상 행위 탐지 방법 워크플로우

Fig. 1. Overall workflow for insider anomaly behavior detection method proposed in this study

3-1 실험 데이터셋

실험에 사용된 데이터는 미국 카네기멜론 대학의 CERT 내부자 위협 센터에서 수집한 Insider Threat Test Dataset r4.2로 가상의 기업에 근무하는 1000명의 내부자에 대한 정보가 담겨 있다[15]. 해당 데이터에 따르면 표 1의 세 가지 시나리오에 따라 기업에 피해를 줄 수 있는 이상 행위를 하는 내부자는 70명이다. 또한 logon.csv, http.csv, email.csv, file.csv, device.csv 파일에 로그온 및 로그오프, 웹사이트 접속, 이메일 전송, 파일 복사, 이동식 드라이브 연결 및 해제 등의 내부자 행위가 각각 기록되어 있으며, 이는 표 2에 정리하였다.

표 1. 내부자에 의해 발생하는 악의적인 행위를 나타내는 시나리오

Table 1. Scenarios of malicious behavior performed by insiders in experimental dataset

Scenario	Contents
Scenario 1	User who did not previously use removable drives or work after hours begins logging in after hours, using a removable drive, and uploading data to wikileaks.org. Leaves the organization shortly thereafter.
Scenario 2	User begins surfing job websites and soliciting employment from a competitor. Before leaving the company, they use a thumb drive (at markedly higher rates than their previous activity) to steal data.
Scenario 3	System administrator becomes disgruntled. Downloads a keylogger and uses a thumb drive to transfer it to his supervisor's machine. The next day, he uses the collected keylogs to log in as his supervisor and send out an alarming mass email, causing panic in the organization. He leaves the organization immediately.

표 2. 내부자 행위가 기록된 파일

Table 2. Files recording insider behavior

File	Insider Behavior
logon.csv	logon, logoff
http.csv	access a web page
email.csv	send a email
file.csv	copy a file
device.csv	connect, disconnect

표 3. 시퀀스 길이에 따른 정상 및 비정상 행위 샘플의 개수

Table 3. Number of normal and abnormal behavior samples according to sequence length

Sequence Length	Number of Normal Behavior Samples	Number of Abnormal Behavior Samples
20	1,637,682	330
30	1,091,604	212
40	818,610	155

표 4. 실험에 사용된 시퀀스 길이에 따른 훈련, 검증, 테스트 샘플의 개수

Table 4. Number of training, validation, and test samples according to sequence length used in the experiment

Sequence Length	Training Samples	Validation Samples	Test Samples
20	40,806	4,535	704
30	34,847	3,872	624
40	31,731	3,526	544

데이터 전처리를 위해 내부자 행위를 숫자로 치환하고, logon.csv, http.csv, email.csv, file.csv, device.csv 파일의 데이터를 결합하여 시간 순서대로 정렬하였다. 이후 데이

터를 순서에 따른 일정 길이의 시퀀스로 분할하여 하나의 샘플로 구성하는 작업을 수행하였다. 시퀀스 길이별 정상 행위 샘플 수와 비정상 행위 샘플 수는 표 3에 기술하였다. 각 시퀀스 길이에 대응하는 샘플을 오토인코더에 입력하기 위해 중복된 것을 상당 부분 제거하고, 원-핫 인코딩을 하여 임베딩 벡터를 생성하였다. 마지막으로 전처리된 데이터를 훈련 데이터(Training Data), 검증 데이터(Validation Data), 테스트 데이터(Test Data)로 나누었다. 정상 행위만 하는 930명의 내부자에 해당하는 샘플을 9:1로 나눠 훈련 데이터와 검증 데이터로 사용하였고, 정상 행위와 비정상 행위를 번갈아 하는 70명의 내부자에 해당하는 샘플을 테스트 데이터로 사용하였다. <표 4>는 실험에 사용된 각 시퀀스 길이에 따른 훈련 데이터, 검증 데이터, 테스트 데이터의 샘플 수를 나타낸다.

3-2 LSTM 기반 잡음 제거 오토인코더(LSTM-DAE) 모델

시계열 기반의 Insider Threat Test Dataset r4.2 데이터를 전처리하여 얻은 정상 행위 샘플만을 사용하여 LSTM 기반 잡음 제거 오토인코더(LSTM-DAE) 모델을 개발하였다. 해당 모델은 입력층(Input Layer), 드롭아웃(Dropout)층 및 LSTM층을 포함하는 인코더와 LSTM층 및 출력층(Output Layer)을 포함하는 디코더로 구성된다. 또한 가변 길이의 시퀀스를 고정 길이의 시퀀스로 변환하여 얻은 샘플을 유용한 특성을 지닌 벡터로 압축하는 과정에서 내재되어 있는 패턴을 학습하는 데 사용된다.

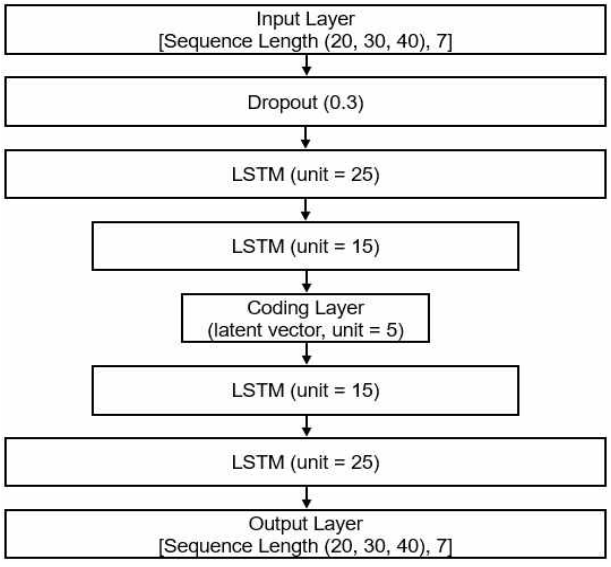

그림 2. 제안하는 LSTM-DAE 모델
Fig. 2. Proposed LSTM-DAE Model

LSTM-DAE 모델의 아키텍처는 그림 2와 같다. 해당 모델의 인코더에는 드롭아웃층을 두어 입력 데이터의 30%를 드롭아웃하여 과대 적합을 방지하고, 차원이 25, 15, 5로 점차 감소하는 LSTM층을 거치면 잠재 벡터가 추출되도록 구성하였다. 디코더에는 잠재 벡터가 차원이 5, 15, 25로 점차 증가하는 LSTM층을 거치면 입력 데이터와 유사한 재구성 데이터가 생성되도록 구성하였다. 또한 모델 컴파일(Compile) 단계에서는 'Adam'을 최적화 알고리즘(Optimizer)으로 사용하였고, 'Mean Squared Error'를 손실 함수(Loss Function)로 사용하였다.

3-3 비지도 기계 학습 기반 LOF와 IF

이상 탐지는 데이터셋의 정상적인 형태가 아닌 의심스러운 사건이나 항목을 찾는 것을 목표로 하며, 머신러닝 분야에서 상당한 관심을 불러일으키고 있다[16]. 본 연구에 사용된 이상 탐지 알고리즘은 비지도 학습을 기반으로 하는 LOF와 IF이다. LOF 알고리즘은 주어진 데이터셋에 속한 각 개체의 제한된 주변만 고려하여 개체가 얼마나 외곽에 있는지 정량화하는 밀도 기반 이상 탐지 알고리즘이다[4]. 이 알고리즘은 분포가 고르지 않은 데이터셋의 이상치 검출에 적합하지만, 계산 비용이 너무 커서 대규모 고차원 데이터셋의 이상치 검출에 사용하기 어렵다[17]. IF 알고리즘은 다른 알고리즘에 비해 상대적으로 더 세세한 범위까지 다루고, 선형 시간 복잡도(Linear Time Complexity)를 가져 낮은 메모리 용량을 요구하는 앙상블 기반 이상 탐지 알고리즘이다[5].

3-4 제안한 내부자 이상 행위 탐지 방법의 성능 평가 지표

일반적으로 비지도 학습 기반 이상 탐지는 차원 축소 과정을 통해 입력 데이터에서 유용한 특성을 추출하고, 이를 이상 탐지 알고리즘에 적용하는 방식으로 진행된다. 본 연구에서는 시계열 기반의 Insider Threat Test Dataset r4.2 데이터를 임베딩 벡터로 변환하고, 3.2절에서 제안한 LSTM-DAE 모델에 입력하여 잠재 벡터를 추출하였다. 이후 추출된 잠재 벡터를 비지도 기계 학습 기반 이상 탐지 알고리즘인 LOF와 IF에 입력하여 내부자 이상 행위 탐지 방법의 성능을 평가하였다. 이를 위해 정확도(Accuracy), 정밀도(Precision), 재현율(Recall) 및 F1-Score 등의 지표를 사용하여 모델의 실효성을 검증하였다.

Ⅳ. 실험 결과 및 분석

표 5. 실험 환경
Table 5. Experimental environment

OS	Ubuntu 18.04.6 LTS
CPU	Intel(R) Xeon(R) Gold 5120
GPU	NVIDIA RTX A5000
RAM	264GB
Python	3.10.9
Scikit-Learn	1.2.1
Keras	2.11.0

표 6. 잠재 벡터의 차원이 5일 때 시퀀스 길이, 정상 행위 샘플 개수의 변화에 따른 내부자 이상 행위 탐지 성능 평가
Table 6. Performance of insider anomaly behavior detection with variation of sequence length and number of normal behavior samples when the coding vector dimension is 5

	Sequence Length	Test Data		Accuracy	Precision	Recall	F1-Score
		Number of Abnormal Behavior Samples	Number of Normal Behavior Samples				
Local Outlier Factor	20	88	88	0.8693	**0.8095**	0.9659	0.8808
		88	176	0.8636	0.7167	**0.9773**	0.8269
		88	264	0.8239	0.5890	**0.9773**	0.7350
	30	78	78	0.9103	**0.8721**	**0.9615**	0.9146
		78	156	0.8718	0.7400	0.9487	0.8315
		78	234	0.8654	0.6607	0.9487	0.7789
	40	68	68	0.8603	**0.8356**	0.8971	0.8652
		68	136	0.8088	0.6465	0.9412	0.7665
		68	204	0.8125	0.5739	**0.9706**	0.7213
Isolation Forest	20	88	88	0.8977	**0.8365**	**0.9886**	0.9062
		88	176	0.8826	0.7568	0.9545	0.8442
		88	264	0.8494	0.6277	0.9773	0.7644
	30	78	78	0.8526	**0.7835**	**0.9744**	0.8686
		78	156	0.8718	0.7308	**0.9744**	0.8352
		78	234	0.8846	0.6944	0.9615	0.8065
	40	68	68	0.8309	**0.7922**	0.8971	0.8414
		68	136	0.8824	0.7500	**0.9706**	0.8462
		68	204	0.8787	0.6882	0.9412	0.7950

표 7. 잠재 벡터의 차원이 7일 때 시퀀스 길이, 정상 행위 샘플 개수의 변화에 따른 내부자 이상 행위 탐지 성능 평가
Table 7. Performance of insider anomaly behavior detection with variation of sequence length and number of normal behavior samples when the coding vector dimension is 7

	Sequence Length	Test Data		Accuracy	Precision	Recall	F1-Score
		Number of Abnormal Behavior Samples	Number of Normal Behavior Samples				
Local Outlier Factor	20	88	88	0.8295	**0.7636**	0.9545	0.8485
		88	176	0.8447	0.6911	**0.9659**	0.8057
		88	264	0.8466	0.6250	**0.9659**	0.7589
	30	78	78	0.8718	**0.8085**	**0.9744**	0.8837
		78	156	0.8675	0.7282	0.9615	0.8287
		78	234	0.8365	0.6098	0.9615	0.7463
	40	68	68	0.8971	**0.8553**	**0.9559**	0.9028
		68	136	0.8676	0.7356	0.9412	0.8258
		68	204	0.8603	0.6500	**0.9559**	0.7738
Isolation Forest	20	88	88	0.8920	**0.8286**	**0.9886**	0.9016
		88	176	0.8409	0.6825	0.9773	0.8037
		88	264	0.8182	0.5800	**0.9886**	0.7311
	30	78	78	0.8590	**0.7857**	**0.9872**	0.8750
		78	156	0.8675	0.7196	**0.9872**	0.8324
		78	234	0.8718	0.6638	**0.9872**	0.7938
	40	68	68	0.8529	**0.7857**	0.9706	0.8684
		68	136	0.8480	0.6907	**0.9853**	0.8121
		68	204	0.8382	0.6091	**0.9853**	0.7528

본 연구에서 제안하는 LSTM-DAE 모델 및 이상 탐지 알고리즘인 LOF와 IF를 사용하여 내부자 이상 행위 탐지 실험 및 성능 평가를 수행하였다. 수행된 실험 환경은 표 5와 같다. LSTM-DAE의 인코더를 사용하여 5차원의 잠재 벡터(인코딩 벡터)를 추출하였고, 추출된 잠재 벡터를 LOF 와 IF 알고리즘에 주입하여 학습시켰다. 정상 행위 샘플들로만 구성된 훈련 데이터 및 검증 데이터와 달리 테스트 데이터는 정상 행위 샘플과 비정상 행위 샘플이 섞여 있어 이를 고려하여 실험을 진행하였다. LOF와 IF에서 우수한 성능을 도출하는 하이퍼파라미터(Hyperparameter)를 얻기 위해 비정상 행위 샘플을 절반으로 나눠 2개의 테스트 데이터셋을 만들고, 이를 LSTM-DAE 모델에 입력하여 5차원인 잠재 벡터를 추출하였다. 테스트 데이터로부터 추출한 2개의 잠재 벡터 중 1개를 LOF와 IF에 입력하여 최적의 하이퍼파라미터를 얻고, 남은 1개의 잠재 벡터를 LOF와 IF에 입력할 때 이전에 얻은 하이퍼파라미터를 사용하였다. 또한 테스트 데이터의 정상 행위 샘플 수를 3배까지 증가시켜 앞선 실험을 반복 수행한 후, 내부자 이상 행위 탐지 방법의 성능을 평가하여 그 결과를 표 6에 정리하였다. 이뿐만 아니라 잠재 벡터의 차원을 5에서 7로 증가시켜 앞선 실험을 다시 반복 수행한 후, 내부자 이상 행위 탐지 방법의 성능을 평가하여 그 결과를 표 7에 정리하였다.

4-1 재현율 분석

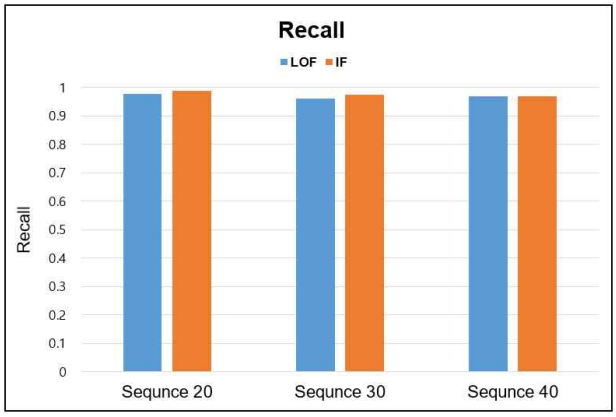

그림 3. 잠재 벡터의 차원이 5일 때 시퀀스 길이에 따른 가장 높은 재현율

Fig. 3. Highest recall according to sequence length when the coding vector dimension is 5

기업 내부 자료, 기술 등의 유출을 방지하려면 이상 행위를 놓치지 않고 탐지하는 것이 중요하므로 본 연구에서는 실제 이상 행위 중 모델이 이상 행위로 판단하는 비율인 재현율을 높이는 데 주력하였으며, 이는 표 6, 표 7 및 그림 3에서 확인할 수 있다. 또한 5차원인 잠재 벡터를 LOF에 입력하면 시퀀스 길이가 20일 때 재현율이 0.9773으로 가장 높았고, IF에 입력하면 시퀀스 길이가 20일 때 재현율이 0.9886으로 가장 높았다. 즉 이상 탐지 알고리즘의 종류와 상관없이 시퀀스 길이가 짧을수록 재현율이 높았고, 이러한 결과가 나타난 이유는 긴 시퀀스에 비해 짧은 시퀀스가 나타낼 수 있는 패턴의 다양성이 적어 모델이 비정상 행위를 놓치지 않고 잘 감지할 수 있었기 때문인 것으로 보인다.

간결성을 위해 7차원인 잠재 벡터를 사용한 실험 결과에 대한 그래프를 첨부하지는 않았지만, 표 7을 살펴보면 재현율이 시퀀스 길이에 상관없이 전반적으로 고르고 높게 분포하는 것을 확인할 수 있다. 또한 5차원인 잠재 벡터를 사용했을 때 가장 낮은 재현율은 0.8971이고, 7차원인 잠재 벡터를 사용했을 때 가장 낮은 재현율은 0.9412인 것을 통해 잠재 벡터의 차원이 증가하면 재현율의 저점도 같이 증가하는 것을 확인할 수 있다. 이는 잠재 벡터의 차원이 증가하여 입력 데이터에 내재된 패턴이 더욱 잘 표현되었기 때문인 것으로 분석된다.

4-2 정밀도 분석

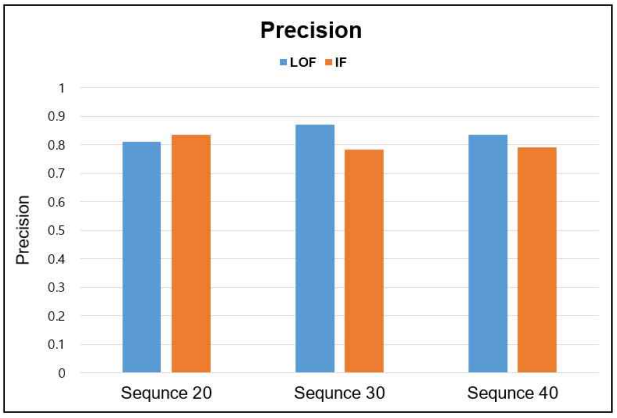

그림 4. 잠재 벡터의 차원이 5일 때 시퀀스 길이에 따른 가장 높은 정밀도

Fig. 4. Highest precision according to sequence length when the coding vector dimension is 5

모델이 이상 행위로 판단한 것 중 실제로 이상 행위인 것의 비율인 정밀도의 경우, 5차원인 잠재 벡터를 LOF에 입력하면 시퀀스 길이가 30일 때 0.8721로 가장 높았고, IF에 입력하면 시퀀스 길이가 20일 때 0.8365로 가장 높았다. 또한 7차원인 잠재 벡터를 LOF에 입력하면 시퀀스 길이가 40일 때 0.8553으로 가장 높았고, IF에 입력하면 시퀀스 길이가 20일 때 0.8286으로 가장 높았다. LOF 사용 시에는 시퀀스 길이가 30 혹은 40일 때 정밀도가 가장 높았지만, <표 6>을 살펴보면 전반적으로 시퀀스 길이가 짧을 때 정밀도가 높은 것을 확인할 수 있다. 이는 재현율과 비슷하게 긴 시퀀스에 비해 짧은 시퀀스가 나타낼 수 있는 패턴의 가짓수가 적어 모델이 비정상 행위를 더 쉽게 구분할 수 있었기 때문인 것으로 보인다.

실제 환경에서는 정상 행위가 비정상 행위보다 더 많이 발생하기 때문에 정상 행위 샘플 수를 증가시킨 테스트 데이터를 이용하여 모델의 성능을 평가할 필요가 있어 비정상 행위 샘플 수를 일정하게 두고, 정상 행위 샘플 수를 3배까지 증가

시켜 실험을 수행하였다. 실험 결과 정상 행위 샘플 수가 증가하면 정밀도가 하락하는 것을 확인할 수 있었다. 이는 이상 탐지 알고리즘이 새로 발생한 정상 행위 샘플을 모두 정상으로 판단하지 못하고 일부를 비정상으로 판단하여 나타나는 현상으로 분석된다.

문헌 [11]에서는 전체 업무 패턴을 통해 얻은 loss 값이 임계값(Threshold)보다 크면 데이터 유출 징후가 있는 것으로 판단하였고, 데이터 유출 징후를 결정하는 데 가장 중요한 임계값을 저자의 주관적인 판단에 따라 결정하였다. 하지만 본 연구에서는 유용한 시퀀스 정보를 담고 있는 잠재 벡터를 추출한 후, 이를 이상 탐지 알고리즘인 LOF와 IF에 입력하여 내부자 이상 행위를 탐지하였다. 그 결과 성능 평가 지표 중 재현율의 경우 문헌 [11]과 본 연구가 비슷한 성능을 보였지만 정밀도의 경우 문헌 [11]보다 본 연구가 더 뛰어난 성능을 보였다. 하지만 전반적으로 정밀도가 재현율에 미치지 못하는데, 모델이 실제 환경에서 유효하려면 정밀도의 성능을 높일 필요가 있어 향후 정밀도의 성능을 개선할 예정이다.

V. 결 론

본 연구에서는 LSTM 기반 잡음제거 오토인코더(LSTM-DAE) 모델을 활용하여 내부자 행위 정보를 기록한 로그 데이터를 분석하고 내부자 이상 행위를 탐지하는 방법을 제안하였다. LSTM-DAE 모델로 생성한 잠재 벡터를 사용하여 비지도 학습 기반 이상 탐지 알고리즘인 LOF와 IF를 학습시키고, 제안한 모델의 성능을 분석하였다. 분석 결과 재현율 지표에서 뛰어난 성능을 보였지만 정밀도 지표에서 실제 환경에 적용하기 어려운 성능을 보였는데, 이는 비지도 학습의 한계로 인해 정밀도 성능 개선에 어려움을 겪었기 때문이다. 따라서 향후에는 임베딩 벡터 차원, 신경망 모델의 층 개수, 활성화 함수, 손실 함수 등 다양한 하이퍼파라미터를 조정하여 재현율에 영향을 주지 않으면서 정밀도를 향상시켜 모델의 성능을 개선할 계획이다.

또한 향후 연구에서는 머신러닝을 기반으로 하는 이상 탐지 알고리즘인 LOF와 IF 외에도 신경망을 기반으로 하는 비지도 학습 이상 탐지 알고리즘인 Deep SVDD, OC-NN (One-Class Neural Network) 등을 활용하여 내부자 이상 행위 탐지 성능을 개선하는 연구를 진행할 것이다. 이뿐만 아니라 내부자의 직위, 부서 등을 추가로 고려한 내부자 이상 행위 탐지 연구를 진행하여 모델의 실효성을 높이기 위한 추가 연구가 필요하다.

감사의 글

본 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2019R1F1A1059036).

참고문헌

[1] S. C. Yoo, A Comparative Evaluation of Autoencoder Anomaly Detection Algorithms, Master's Thesis, Ajou University, Suwon, February 2023.

[2] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System," *IEEE Access*, Vol. 7, pp. 41525-41550, April 2019. https://doi.org/10.1109/ACCESS.2019.2895334

[3] S. Yuan and X. Wu, "Deep Learning for Insider Threat Detection: Review, Challenges and Opportunities," *Computers & Security*, Vol. 104, 102221, May 2021. https://doi.org/10.1016/j.cose.2021.102221

[4] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "LOF: Identifying Density-Based Local Outliers," in *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data (SIGMOD '00)*, Dallas: TX, pp. 93-104, May 2000. https://doi.org/10.1145/342009.335388

[5] F. T. Liu, K. M. Ting, and Z. H. Zhou, "Isolation Forest," in *Proceedings of the 2008 8th IEEE International Conference on Data Mining*, Pisa, Italy, pp. 413-422, December 2008. https://doi.org/10.1109/ICDM.2008.17

[6] B. Schölkopf, R. C. Williamson, A. Smola, J. Shawe-Taylor, and J. Platt, "Support Vector Method for Novelty Detection," in *Proceedings of the 12th International Conference on Neural Information Processing Systems (NIPS '99)*, Denver: CO, pp. 582-588, November 1999.

[7] G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," *Science*, Vol. 313, No. 5786, pp. 504-507, July 2006. https://doi.org/10.1126/science.1127647

[8] Z. Zhang and X. Deng, "Anomaly Detection using Improved Deep SVDD Model with Data Structure Preservation," *Pattern Recognition Letters*, Vol. 148, pp. 1-6, August 2021. https://doi.org/10.1016/j.patrec.2021.04.020

[9] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 2nd ed. Sebastopol, CA: O'Reilly Media, 2019.

[10] Y.-J. So, H.-S. Ji, and K.-C. Kang, "Analysis of Global Research Trends and Utilized Datasets for Anomaly Detection," in *Proceedings of 2022 Korea Software Congress*, Jeju, pp. 1718-1720, December 2022.

[11] M. Sakurada and T. Yairi, "Anomaly Detection Using

Autoencoders with Nonlinear Dimensionality Reduction," in *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis (MLSDA '14)*, Gold Coast, Australia, pp. 4-11, December 2014. https://doi.org/10.1145/2689746.2689747

[12] S.-J. Kim and T.-S. Shon, "LSTM Autoencoder-Based Insider Data Leak Detection," *Journal of Digital Contents Society*, Vol. 23, No. 6, pp. 1159-1166, June 2022. https://doi.org/10.9728/dcs.2022.23.6.1159

[13] E. Pantelidis, G. Bendiab, S. Shiaeles, and N. Kolokotronis, "Insider Threat Detection using Deep Autoencoder and Variational Autoencoder Neural Networks," in *Proceedings of 2021 IEEE International Conference on Cyber Security and Resilience (CSR)*, Rhodes, Greece, pp. 129-134, July 2021. https://doi.org/10.1109/CSR51186.2021.9527925

[14] V. Koutsouvelis, S. Shiaeles, B. Ghita, and G. Bendiab, "Detection of Insider Threats using Artificial Intelligence and Visualisation," in *Proceedings of 2020 6th IEEE Conference on Network Softwarization (NetSoft)*, Ghent, Belgium, pp. 437-443, June 2020. https://doi.org/10.1109/NetSoft48620.2020.9165337

[15] Carnegie Mellon University. Insider Treat Test Dataset [Internet]. Available: https://kilthub.cmu.edu/articles/dataset/Insider_Threat_Test_Dataset/12841247/1/.

[16] O. Alghushairy, R. Alsini, T. Soule, and X. Ma, "A Review of Local Outlier Factor Algorithms for Outlier Detection in Big Data Streams," *Big Data and Cognitive Computing*, Vol. 5, No. 1, 1 December 2020. https://doi.org/10.3390/bdcc5010001

[17] Z. Cheng, C. Zou, and J. Dong, "Outlier Detection using Isolation Forest and Local Outlier Factor," in *Proceedings of the Conference on Research in Adaptive and Convergent Systems (RACS '19)*, Chongqing, China, pp. 161-168, September 2019. https://doi.org/10.1145/3338840.3355641

심기천(Ki-Chun Sim)

2021년 : 아주대학교 수학과(학사)
2023년 : 아주대학교 일반대학원 지식정보공학과(석사과정)

2021년～현 재: 아주대학교 대학원 지식정보공학과 석사과정
※관심분야 : 기계학습(Machine Learning), 정보보호(Information Security), 블록체인(Blockchain)

김강석(Kangseok Kim)

2007년 : Indiana University (at Bloomington) 컴퓨터공학과 (공학박사)

2010년～2016년: 아주대학교 대학원 지식정보공학과 연구교수
2016년～현 재: 아주대학교 사이버보안학과 부교수
※관심분야 : 정보보안(Information Security), 딥러닝 응용보안(Applied Deep Learning for Security)