Journal of Convergence for Information Technology
Vol. 12. No. 2, pp. 47-53, 2022
e-ISSN 2586-4440
DOI : https://doi.org/10.22156/CS4SMB.2022.12.02.047

이상 탐지를 위한 시스템콜 시퀀스 임베딩 접근 방식 비교

이근섭[1], 박경선[1], 김강석[2*]
[1]아주대학교 대학원 지식정보공학과 학생, [2]아주대학교 사이버보안학과 교수

Comparison of System Call Sequence Embedding Approaches for Anomaly Detection

Keun-Seop Lee[1], Kyungseon Park[1], Kangseok Kim[2*]
[1]Student, Dept. of Knowledge Information Engineering, Graduate School of Ajou University
[2]Professor, Dept. of Cyber Security, Ajou University

요 약 최근 지능화된 보안 패러다임의 변화에 따라, 다양한 정보보안 시스템에서 발생하는 각종 정보를 인공지능 기반 이상탐지에 적용하기 위한 연구가 증가하고 있다. 따라서 본 연구는 로그와 같은 시계열 데이터를 수치형 특성인 벡터로 변환하기 위하여 딥러닝 기반 Word2Vec 모델의 CBOW와 Skip-gram 추론 방식과 동시발생 빈도 기반 통계 방식을 사용하여 공개된 ADFA 시스템콜 데이터에 대하여, 벡터의 차원, 시퀀스 길이 및 윈도우 사이즈를 고려한 다양한 임베딩 벡터로의 변환에 대한 실험을 진행하였다. 또한 임베딩 모델로 생성된 벡터를 입력으로 하는 GRU 기반 이상 탐지 모델을 통해 탐지 성능뿐만 아니라 사용된 임베딩 방법들의 성능을 비교 평가하였다. 통계 모델에 비해 추론 기반 모델인 Skip-gram이 특정 윈도우 사이즈나 시퀀스 길이에 치우침 없이 좀 더 안정되게(stable) 성능을 유지하여, 시퀀스 데이터의 각 이벤트들을 임베딩 벡터로 만드는데 더 효과적임을 확인하였다.

주제어 : 침입탐지시스템, 이상탐지, 시스템콜, 임베딩, GRU

Abstract Recently, with the change of the intelligent security paradigm, study to apply various information generated from various information security systems to AI-based anomaly detection is increasing. Therefore, in this study, in order to convert log-like time series data into a vector, which is a numerical feature, the CBOW and Skip-gram inference methods of deep learning-based Word2Vec model and statistical method based on the coincidence frequency were used to transform the published ADFA system call data. In relation to this, an experiment was carried out through conversion into various embedding vectors considering the dimension of vector, the length of sequence, and the window size. In addition, the performance of the embedding methods used as well as the detection performance were compared and evaluated through GRU-based anomaly detection model using vectors generated by the embedding model as an input. Compared to the statistical model, it was confirmed that the Skip-gram maintains more stable performance without biasing a specific window size or sequence length, and is more effective in making each event of sequence data into an embedding vector.

Key Words : IDS, Anomaly Detection, System Call, Embedding, GRU

*This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT: Ministry of Science and ICT) (No. NRF-2019R1F1A1059036).
*This article is extended and excerpted from Master Thesis.

*Corresponding Author : Kangseok Kim (kangskim@ajou.ac.kr)

Received December 27, 2021 Revised January 28, 2022
Accepted February 20, 2022 Published February 28, 2022

1. 서론

현재 우리는 인터넷을 통해 정보 공유와 사회적 상호작용을 생성하고 강화하며 확장해 나가는 시기에 있다. 하지만, 인터넷을 통한 방대한 정보 및 서비스, 네트워크의 사용은 랜섬웨어와 같은 악성코드 감염, 자동화된 스웜 공격, 딥페이크 기술을 활용한 해킹 등 지속적으로 진화하는 사이버 보안 위협에 노출될 수 있다. 따라서 지능화된 사이버 위협에 대응하기 위한 새로운 보안 기술의 개발이 필요하며, 일반적으로 시스템에서 수집되는 로그와 같은 시계열 데이터는 호스트 및 네트워크에서 발생하는 사용자의 악의적인 활동이나 외부의 공격으로 의심되는 이벤트와 관련된 다양한 정보를 가지고 있어 효율적인 이상 징후 탐지 시스템을 구축하는데 도움이 된다. 기존의 이상 탐지 시스템은 시그니처 및 룰 기반의 시나리오에 기반 하여, 정보 자원에 불법으로 접근하거나 자원을 고갈시키는 이상 행위(anomaly behavior)를 탐지 및 차단하고 있어, 새로운 공격 벡터(unknown attack vectors)의 탐지 및 이에 대처하기에는 한계점이 있다. 이러한 한계점을 극복하기 위하여 다양한 정보보안 시스템에서 발생하는 각종 정보를 인공지능 기반 이상 탐지 모델에 적용함으로써, 오탐 제거와 알려지지 않은 미탐 공격을 탐지할 수 있는 지능형 이상 징후 탐지 기술에 대한 연구가 증가하고 있다. 그러므로, 본 연구는 로그와 같은 시계열 데이터의 벡터화를 위한 임베딩 방법과 추출된 임베딩 벡터를 지도학습으로의 전이학습을 통하여 진화하는 보안 위협을 효과적으로 탐지할 수 있는 이상 징후 탐지 방법을 개발하는데 있다. 따라서 실험에 사용된 시퀀스(시계열) 데이터(ADFA-LD: Australian Defence Force Academy Linux Dataset[1])에 대한 이해를 높이기 위하여, 통계 기반 모델뿐만 아니라 딥러닝 네트워크로 하여금 데이터에 내재된 의미 있는 정보(semantic information)의 도출을 위한 추론 기반 모델을 사용하여 학습시키고, 학습된 정보를 입력으로 하는 순환 신경망 기반의 GRU(Gated Recurrent Units)[2] 모델을 통해서 이상 탐지 정확도를 높일 뿐만 아니라 사용되어진 시퀀스 데이터의 벡터화를 위한 방법들의 성능을 비교하였다.

본 연구의 목차는 2장에서 관련 연구를 기술하고, 3장에서는 제안하는 연구 방법을 설명하며, 4장에서 실험 결과를 분석하고, 5장에서 결론 및 향후 연구에 대하여 논한다.

2. 관련 연구

HIDS(Host-based Intrusion Detection System) 관련연구로써 Xie et al.[3]은 HIDS의 비정상행위 탐지 방식(HADS: anomaly-based HIDS)의 개발을 위해 ADFA 리눅스 데이터셋의 패턴과 빈도를 이용한 분석뿐만 아니라 KNN-based HADS를 ADFA 리눅스 데이터셋을 이용해 평가했다. 일부 공격들에 대해 수긍할 만한 결과를 보였지만 여전히 현재 컴퓨터 시스템의 복잡한 행위들과 다양한 공격들에 대한 연구가 부족하다는 단점이 있다. Creech et al.[4]은 HIDS는 오경보율로 인해 설계가 매우 까다로운데, 이에 대하여 새로운 특징을 도출하기 위해 의미론적 구조를 HIDS의 커널 수준 시스템 콜에 적용하는 이상 탐지 접근 방법론이 연구 되었다. Maske et al.[5]은 ADFA 데이터의 시스템콜 트레이싱을 사용하여, 전처리로 단어 사전을 구성하고 이 단어 사전을 통해 구문 사전을 생성하여 특성 벡터를 추출하고, 이에 대한 콜 트레이싱 추론 방법을 사용하는 의미론적 해석을 통해 탐지하는 방법을 제안하였다. Aghaei[6]는 앙상블(Ensemble) 분류를 이용한 빈도 기반 오용 탐지 방식을 개발하였다. N-gram을 이용해 전처리 후, 특징들을 추출하여 패턴을 생성하였고 SMOTE 알고리즘을 통해 클래스별 패턴 수의 균형을 맞췄다. 분류 모델은 SVM, PART, decision tree, random forest의 다수 투표 방식 앙상블 기법을 기반으로 하였고 제안된 오용 침입 탐지 방식은 공격탐지에 좋은 성능을 보였다. Borisaniya et al.[7]은 시스템콜 트레이싱을 이용해 진행 중인 프로세스의 비정상 행위를 탐지하는 것은 기계 학습을 통해 다뤄지는 전형적인 패턴의 인지 문제라고 판단하여 패턴을 추출하기 위한 기계학습의 다양한 분류 알고리즘을 이용하였다. ADFA 리눅스 데이터에 대한 향상된 벡터 공간 표현 기법의 성능을 평가하였고 결과적으로 시스템콜을 통해 프로세스의 행위를 구별하는 데에 좋은 성능을 보였다. Kwon et al.[8]은 합성곱 신경망(Convolutional Neural Networks) 모델과 NSL-KDD[9] 데이터 세트를 사용하여 이상 탐지 모델을 개발했으며, 모델 깊이에 따라 탐지 성능이 증가하지 않는다는 것을 보였다. Fu et al.[10]은 LSTM(Long Short-Term Memory) 기반 이진분류 모델을 제안했

으며, NSL-KDD 데이터셋을 기반으로 성능을 평가한 결과 탐지율과 분류 정확도 면에서 우수한 분류 성능을 보였다. Kim et al.[11]은 시퀀스 데이터가 RNN-AE(Recurrent Neural Network-based Auto-Encoder) 와 RNN-DAE(Recurrent Neural Network-based Denoising Auto-Encoder)를 사용하면서 고정길이 벡터로 변환되어지고, 변환된 벡터들이 이상 탐지 모델을 학습시키기 위하여 사용되는 메커니즘으로 구성되는 탐지 방법론을 기술하였다.

본 연구에서는 ADFA 시스템콜 데이터에 대하여, 벡터의 차원, 시퀀스 길이 및 윈도우 사이즈를 고려한 다양한 임베딩 벡터로의 변환에 대한 실험을 진행하였다. 또한 임베딩 모델로 생성된 벡터를 입력으로 하는 GRU 기반 이상 탐지 모델을 통해 탐지 성능뿐만 아니라 사용된 임베딩 방법들의 성능을 비교 평가하였다.

3. 제안 연구방법

사용된 데이터의 구성을 간략히 살펴보고 데이터 전처리 후, 훈련 데이터에 통계와 추론 기반 임베딩 기법을 적용하여 생성된 임베딩벡터를, GRU 분류모델에 입력 데이터로 주입하여 학습시키고, 테스트 데이터로 학습된 모델의 분류 성능을 평가하였다. Fig. 1은 본 연구에서 제안하는 연구 방법의 흐름도(Workflow)이다.

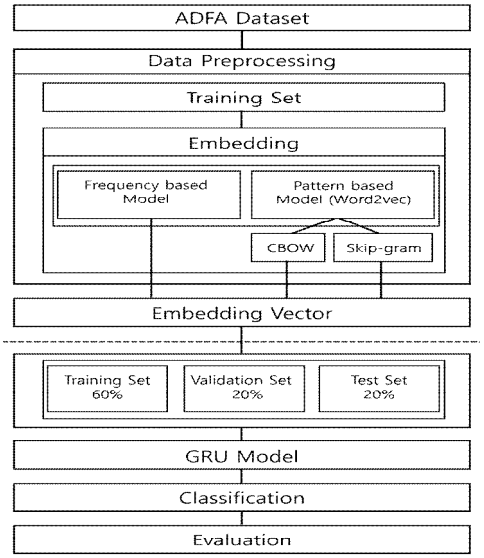

Fig. 1. A Simplified Workflow for Proposed Anomaly Detection Methodology

3.1 데이터 셋

사용된 공개 ADFA-LD [1] 데이터는 리눅스 로컬 서버로부터 수집된 다양한 애플리케이션에서 발생한 가장 최근의 공격 및 취약점에 대한 system call trace로 구성되어 있다. Table 1에서와 같이 데이터셋은 세 개의 데이터 그룹으로 구성되어 있고 각 그룹에는 system call trace 파일들이 존재한다. 정상(normal) 데이터인 TDM(Training Data Master)과 VDM(Validation Data Master)이 있으며, 공격(attack) 데이터인 ADM(Attack Data Master)으로 구성되어 있다. ADM은 'Adduser'와 'Hydra-FTP', 'Hydra-SSH', 'JavaMeterpreter', 'Meterpreter', 'Web-Shell'의 6가지 공격 데이터로 구성되어 있다.

Table 1. Data Groups in the ADFA-LD Dataset

Data group	Type of traces	Number of traces
TDM	Normal	833
VDM	Normal	4372
ADM	Adduser	91
	Hydra-FTP	162
	Hydra-SSH	176
	JavaMeterpreter	124
	Meterpreter	75
	Web-Shell	118

3.2 임베딩 벡터 모델(Embedding Vector Model)

프로그램의 실행에 관한 정보를 기록한 로그(e.g. 시스템 콜 트레이싱 등)와 같은 시계열 데이터를 수치형 특성인 벡터로 변환하기 위하여 빈도 기반 통계 방식과 패턴 기반 추론 방식의 임베딩 모델을 사용하였다.

빈도 기반 통계 모델은 시스템콜 이벤트 앞뒤로 다른 시스템콜 이벤트들과 동시에 얼마나 자주 나타나는지 그 빈도를 측정하여 시스템콜 시퀀스를 임베딩하는 방식이다. 시스템콜들로 구성된 행렬을 만들고 윈도우 길이에 따른 각 시스템콜 이벤트의 동시 발생 횟수를 카운팅하여 행렬의 각 요소를 전체 시스템콜 이벤트 동시발생 빈도의 합으로 나누어 시스템콜 이벤트들에 대한 동시 발생 확률분포 형태로 변환하였다. 또한 변환된 값에 로그를 취해주고 음수(-)를 취하여 값을 양수로 바꾸어 주었다. (아래식에서 w_{ij}는 정규화 이전 matrix의 성분)

정규화한 matrix의 성분 $W_{ij} = -\log_2(\frac{w_{ij}}{\sum w_{ij}})$

데이터 셋에 있는 시스템콜의 갯수는 총 175개로 이를 벡터화(vectorization)하게 되면 최대 175(행)x175(열) = 30,625 원소를 가지는 vector matrix가 생성이 되는데 이를 그대로 시스템콜 시퀀스에 적용하게 되면 sparse matrix가 생성되어 각 시스템콜 간의 유사성을 파악하기가 어렵다. 따라서 유의미한 벡터들로 차원 축소를 하기 위하여 Truncated SVD(Truncated Singular Value Decomposition)를 사용하여 차원 축소 후 밀집 벡터 형태로 변환하였다. 본 실험에서는 Truncated SVD를 사용하여 Vector Dimension을 [4, 6, 8, 10, 12, 15]로 최소 4에서 최대 15까지 차원을 축소하여 실험하였다. 추론 기반 임베딩 모델을 사용하기 위하여, 자연어처리에서 단어의 의미를 컴퓨터가 이해할 수 있도록 벡터화 하는데 많이 사용되어지는 Word2Vec[12, 13] 모델을 사용하였다. Word2Vec은 CBOW(Continuous Bag of Words) 방식과 Skip-gram 방식이 있다. CBOW는 주변에 있는 단어(맥락)들을 통해 중간에 있는 단어(타킷 단어)들을 예측하는 방법이며, Skip-gram은 단어(타킷 단어)를 기준으로 주변의 단어가 어떤 단어가 올지 예측하는 방법이다. 따라서 로그의 각 시스템콜 이벤트에 대응되는 벡터 생성을 위하여, 비지도 학습 딥러닝 네트워크(2층 완전 연결 신경망)를 사용하여 학습 후 실험에 필요한 특성 벡터들을 추출하였다.

3.3 GRU 기반 이상 탐지 모델

이상 탐지 분류를 위해 GRU 모델을 사용하였으며 앞에서 언급된 임베딩 방식들을 사용하여 시스템콜 이벤트들을 벡터화하여 GRU 기반 이상 탐지 모델에 주입하였다. 임베딩 벡터를 만들기 위하여 윈도우 길이, 생성된 벡터의 원소 수 (차원)을 달리하여 모델을 생성하였다. 시퀀스 GRU 모델 노드의 개수는 시스템콜 시퀀스 길이에 따라 결정된다. 과대적합을 방지하기 위해서 드롭아웃(dropout) 비율을 0.5로 하여 규제를 하였다. 또한 과대적합은 피하면서 모델의 표현력을 높이기 위해서 층을 2개 더 쌓아서 실험하였다. Dense layer에서 활성화(activation) 함수로는 Relu를 사용하였다. Fig. 2가 본 연구에서 사용된 GRU 기반 이상 탐지 모델이다.

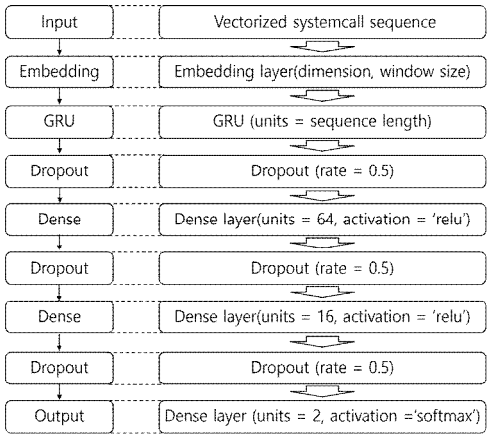

Fig. 2. GRU based Anomaly Detection Model

4. 실험방법 및 결과

4.1 실험 방법

본 실험에서는 전체 데이터 세트를 훈련세트 60%, 검증세트 20%, 테스트세트 20%의 비율로 나누어서 실험을 진행하였다. 또한 통계 및 추론 기반 임베딩 방법이 이상 탐지 성능에 어떻게 영향을 미치는지 확인하기 위하여 윈도우 사이즈(window size)를 [1, 2, 3]으로 길이를 바꿔가면서, 벡터 원소 수(vector dimension)를 [4, 6, 8, 10, 12, 15]로 최소 4에서 최대 15까지 차원을 축소하여 생성된 임베딩 벡터를 GRU 기반 이상 탐지 모델에 입력 데이터로 주입하여 실험하였다.

GRU 모델의 성능을 평가하기 위하여 batch size는 1024로 에폭(epoch)은 400으로 학습하였으며, 더 이상 성능이 나아지지 않을 경우 훈련을 멈추도록 조기종료(early stop)를 사용하였다. 또한 샘플 시퀀스를 [200, 300, 400]의 길이로 실험을 하였다. 각 벡터 원소 수, 윈도우 길이, 시퀀스 길이 별로 통계 및 추론 기반 기법들의 성능 비교를 위하여 GRU 기반의 이상 탐지 모델의 탐지 성능 평가로 비교 분석하였다.

4.2 결과 및 분석

시퀀스 길이 별 데이터의 개수는 200일 때 훈련과 검증 세트가 정상 9,701개, 공격 1,290개이고 테스트 세트가 정상 2449개, 공격 299개이다. 시퀀스 길이가 늘어날 때 데이터 개수가 시퀀스 길이가 늘어난 만큼 줄어들지만 그 비율을 비슷하게 유지하여 실험을 진행

하였다. 각각의 임베딩 모델을 비교하기 위해서 시퀀스 길이와 윈도우 길이 별로 가장 성능이 좋은 벡터 차원을 찾아보았고, Fig. 3은 결과 그래프이다. 빈도 기반 통계 모델은 특정 윈도우 길이에서 성능이 높게 나왔지만 추론 기반 모델인 skip-gram은 빈도 기반 통계 모델과 CBOW에 비해 특정 윈도우 길이나 시퀀스 길이에 치우침 없이 좀 더 안정(stable)되게 성능을 유지하였다. Fig. 4는 시퀀스 길이 400에서의 F1-score 그래프이다. 통계 기반 모델은 윈도우 길이: 2, 벡터 차원: 15에서 F1-score: 0.87로 가장 성능이 높았으며, CBOW 방식은 윈도우 길이: 1, 벡터 차원: 12에서 F1-score: 0.86으로 가장 성능이 높았고, Skip-gram 방식은 윈도우 길이: 2, 벡터 차원: 6에서 F1-score: 0.86으로 가장 성능이 높았다.

본 실험에서 시스템 콜 데이터 표현을 위한 임베딩 벡터 차원의 8차원 길이까지 탐지 모델의 전체적인 성능을 높이지만, 10차원 이상에서는 성능이 떨어지는 경우도 발생하여 임베딩 벡터 차원을 증가시켜도 성능이 향상되지 않는 것을 확인하였다. 또한 임베딩 벡터 차원이 증가할수록 처리 연산량이 증가하기 때문에 벡터의 적절한 차원 크기를 고려하는 것이 필요하다.

윈도우 길이가 증가함에 따라 주변 시스템 콜 시퀀스(맥락)의 사용이 증가하여 훈련 데이터에 과대적합(overfitting)될 수 있고 일반화 문제로 인해 성능이 저하될 수 있다. 따라서 임베딩 벡터 차원의 크기와 마찬가지로 적절한 길이의 윈도우 사용을 고려할 필요가 있다.시퀀스 길이가 증가함에 따라 시퀀스의 더 많은 시계열 패턴을 유지하기 때문에, 본 실험에서 F1-score가 시퀀스 길이에 따라 향상되는 것처럼 각 샘플의 시퀀스 길이가 성능에 영향을 미칠 수 있다. 그러나 너무 긴 시퀀스는 기울기 소실(gradient descent) 문제를 발생시킬 수 있어 다양한 실험을 통해 적절한 시퀀스를 고려하는 것이 필요하다. 실험에서 사용한 시퀀스 길이보다 더 다양하게 길이를 조정하여 실험 데이터에 대한 적절한 시퀀스 길이를 찾아볼 필요가 있으며, 향후 입력 데이터에 대한 가변 길이를 처리하기 위한 마스킹 기법을 사용하여 입력 데이터를 처리하는 것이 필요하다.

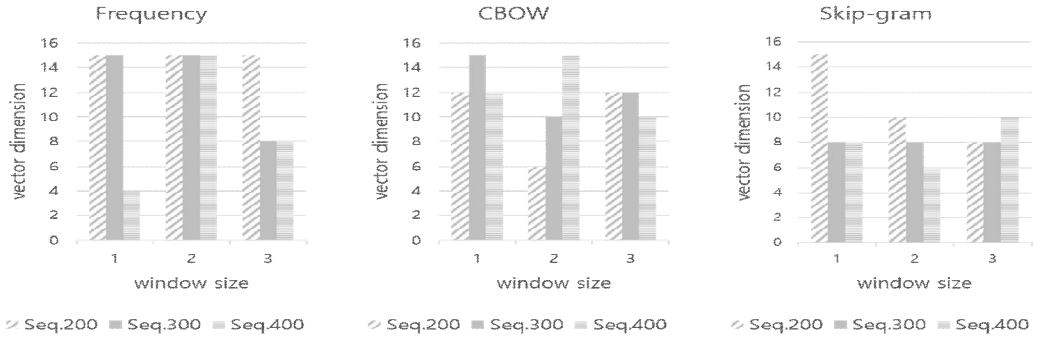

Fig. 3. vector dimension with best score by sequence length and window size

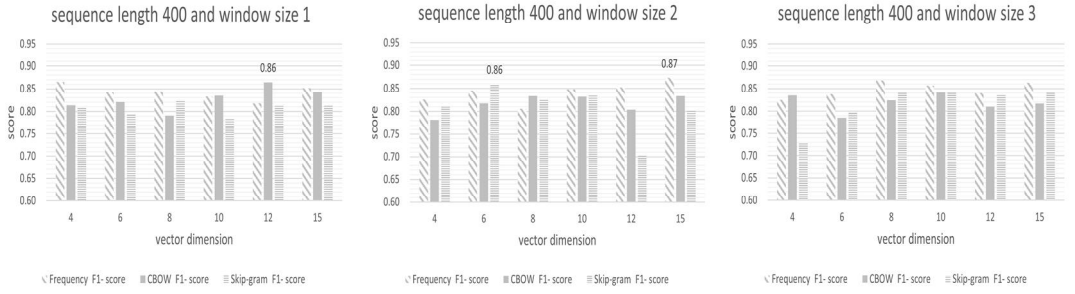

Fig. 4. F1-score for sequence length 400

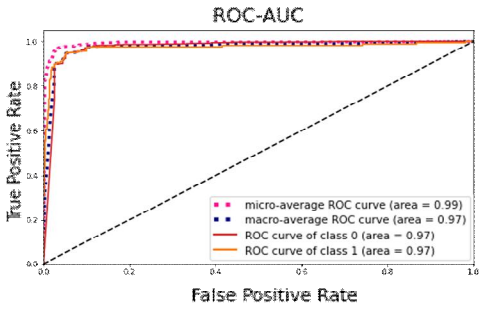

Fig. 5. Experimental Result : ROC-AUC in Skip-gram Model(Window size: 2 and Vector Dimension: 6)

Fig. 5는 Skip-gram 기법으로 생성된 임베딩 벡터(윈도우 길이: 2, 벡터 차원: 6)를 사용하여 학습된 GRU 기반의 이상 탐지 모델의 AUC-ROC 성능 평가 지표를 보인다.

5. 결론

자연어 처리에서 단어의 출현 패턴과 문맥을 통해서 단어들을 정적으로 벡터화 하는 딥러닝 기반 Word2Vec 모델의 CBOW와 Skip-gram 추론 기법과 동시 발생 빈도 기반 통계 기법을 통해서 공개된 ADFA 시스템콜 시퀀스 데이터에 대하여, 벡터의 차원, 데이터 시퀀스 길이 및 윈도우 사이즈를 고려한 다양한 임베딩 벡터로의 변환을 통해 실험을 진행하였다. 또한 임베딩 모델로 생성된 벡터를 입력으로 하는 GRU 기반 이상 탐지 모델을 통해 탐지 성능뿐만 아니라 사용된 임베딩 방식들의 성능을 비교 평가하였다. 빈도 기반 통계 모델에 비해 추론 기반 모델인 Skip-gram이 특정 윈도우 사이즈나 시퀀스 길이에 치우침 없이 좀 더 안정되게(stable) 성능을 유지하여, 시퀀스 데이터의 각 이벤트들을 임베딩 벡터로 만드는데 더 효과적임을 확인하였다.

향후 연구는 특정 모델의 성능이 절대적으로 더 우수한 것이 아니기 때문에 탐지 성능 차이에 영향을 주는 요소는 무엇이고 어떻게 영향을 끼치는가를 살펴보는 것이 더 의미 있을 것이라 생각되어 향후 이 부분에 대한 연구를 보완할 것이다.

또한 동적 문맥 기반 임베딩 방식인 트랜스포머(Transformer)[14] 방식을 활용하여 로그 각각의 이벤트 원소들의 시퀀스 문맥을 고려한 벡터로의 변환을 통해서, 개발된 임베딩 벡터 모델의 자체적인 평가(intrinsic evaluation) 뿐만 아니라 이상 탐지의 성능(extrinsic evaluation)을 향상시키기 위한 연구를 진행할 것이다.

REFERENCES

[1] G. Creech & J. Hu. (2013). Generation of a new IDS test dataset: Time to retire the KDD collection. *IEEE WCNC(Wireless Communications and Networking Conference).*
DOI : 10.1109/WCNC.2013.6555301

[2] K. Cho, B. V. Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk & Yoshua Bengio. (2014). Learning phrase Representations using RNN encoder-decoder for statistical machine translation. *EMNLP,* 1724-1734. *arXiv:1406.1078.*
https://arxiv.org/pdf/1406.1078.pdf

[3] M. Xie & J. Hu. (2013). Evaluating host-based anomaly detection systems: a preliminary analysis of ADFA-LD. *6th IEEE International Congress on Image and Signal Processing (CISP '03),* 1711-1716.
DOI : 10.1109/CISP.2013.6743952

[4] G. Creech, & J. Hu. (2014). A semantic approach to host-based intrusion detection systems using contiguous and discontiguous system call patterns. *IEEE Transactions on Computers, 63(4).*
DOI : 10.1109/TC.2013.13

[5] S. A. Maske & T. J. Parvat. (2016. Aug.). Advanced anomaly intrusion detection technique for host based system using system call patterns. *International Conference on Inventive Computation Technologies (ICICT).* Coimbatore, India.
DOI : 10.1109/INVENTIVE.2016.7824846

[6] E. Aghaei. (2017). Machine learning for host-based misuse and anomaly detection in UNIX environment. *Master Thesis, Computer Science in University of Toledo.*
DOI : 10.13140/RG.2.2.19382.73283

[7] B. Borisaniya & D. Patel. (2015). Evaluation of modified vector space representation using ADFA-LD and ADFA-WD datasets. *Journal of Information Security, 6(3),* 250-264.
DOI : 10.4236/jis.2015.63025

[8] D. Kwon, K. Natarajan, S. C. Suh, H. Kim & J. Kim. (2018. July). An empirical study on network anomaly detection using convolutional neural

networks. *Proceedings of IEEE 38th International Conference Distributed Computing Systems(ICDCS)*, 1595-1598.
DOI: 10.1109/ICDCS.2018.00178

[9] Canadian Institute for Cybersecurit. (n. d.). *NSL-KDD Dataset.* UNB(Online).
https://www.unb.ca/cic/datasets/nsl.html

[10] Y. Fu, F. Lou, F. Meng, Z. Tian, H. Zhang & F. Jiang. (2018. June). An intelligent network attack detection method based on RNN. *Proceedings of IEEE 3rd International Conference Data Science Cyberspace (DSC)*, 483-489.
DOI : 10.1109/DSC.2018.00078

[11] C. Kim, M. Jang, S. Seo, K. Park & P. Kang. (2021). Intrusion detection based on sequential information preserving log embedding methods and anomaly detection algorithms. *IEEE Access, 9,* 58088-58101.
DOI : 10.1109/ACCESS.2021.3071763

[12] T. Mikolov, K. Chen, G. Corrado & J. Dean. (2013). Efficient estimation of word representations in vector space. *ICLR.* arXiv:1301.3781v3.
https://arxiv.org/pdf/1301.3781.pdf

[13] T. Mikolov, I. Sutskever, K. Chen, G. Corrado & J. Dean. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems* (NIPS).
https://papers.nips.cc/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf

[14] A. Vaswan, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser & I. Polosukhin. (2017). Attention is all you need. *31st Conference on Neural Information Processing Systems (NIPS).* arXiv:1706.03762v5

이 근 섭(Keun-Seop Lee) [정회원]

· 2004년 2월 : 서울시립대학교 수학과 (이학사)
· 2020년 3월 ~ 현재 : 아주대학교 지식정보공학과 학생 (석사과정)
· 관심분야 : 사이버보안, 기계학습
· E-Mail : goodinsght@ajou.ac.kr

박 경 선(Kyungseon Park) [정회원]

· 2019년 2월 : 한밭대학교 정보통신공학과 (학사)
· 2020년 3월 ~ 현재 : 아주대학교 지식정보공학과 학생 (석사과정)
· 관심분야 : 이상탐지, 기계학습, 정보보안
· E-Mail : gseon130@ajou.ac.kr

김 강 석(Kangseok Kim) [정회원]

· 2007년 11월 : Indiana University at Bloomington 컴퓨터 공학(박사)
· 2010년 9월 ~ 2016년 2월 : 아주대학교 대학원 지식정보공학과 연구교수
· 2016년 3월 ~ 현재 : 아주대학교 사이버보안학과 부교수
· 관심분야 : 빅데이터 응용보안, 기계학습 및 딥러닝
· E-Mail : kangskim@ajou.ac.kr