중요 특성 기반 저복잡 네트워크 이상탐지 방법 연구

김미르, 권민혜
숭실대학교

48rlaalfm@soongsil.ac.kr, minhae@ssu.ac.kr

Low-complex Anomaly Detection Method Using Important Features

Miru Kim, Minhae Kwon
Soongsil University

요 약

본 논문에서는 저사양 IoT 디바이스에서도 사용가능한 네트워크 이상탐지 방법에 대한 연구를 진행한다. 일반적인 딥러닝 기반의 네트워크 이상탐지 방법은 좋은 성능을 가지지만, 연산 복잡도가 크기 때문에 GPU, NPU 와 같은 고성능의 하드웨어를 요구하게 된다. 따라서 IoT 디바이스와 같이 저사양의 디바이스에서는 딥러닝 기반의 네트워크 이상탐지 시스템의 적용이 어렵다. 이에 본 논문에서는 딥러닝 모델 기반의 이상탐지 방법에서 사용하는 중요 데이터 특성을 파악하고, 이를 이용하여 Isolation Forest 기반의 저복잡도 네트워크 이상탐지 방식을 제안한다. 또한 일반적인 오토인코더 기반의 방식과 제안하는 중요 피처 기반의 Isolation Forest 방식의 성능 비교 실험을 진행한 결과, 제안하는 방식의 F1-score 가 오토인코더 기반의 방식의 93%를 달성하였음을 확인하였다.

Ⅰ. 서 론

인공신경망 구조를 사용하는 딥러닝 기술이 발전됨에 따라 많은 산업 분야에서 이를 적극적으로 활용하고자 하고 있다. 예를 들어 네트워크 이상탐지 상황과 같이 네트워크 침입 시도가 급변하는 경우, 정상적인 네트워크 트래픽만을 이용하여 침입 시도 탐지를 진행하기 위해 딥러닝 모델인 오토인코더가 활용될 수 있다[1]. 하지만 오토인코더 모델은 복잡한 은닉층의 구조로 인하여 연산 복잡도가 크기 때문에 저사양 IoT 디바이스와 같은 네트워크 단말에서의 적용이 어렵다.

따라서 본 논문에서는 저사양 IoT 디바이스에서도 사용 가능한 저복잡도 네트워크 이상탐지 방법을 제안한다. 제안하는 방법은 데이터 특성 교체 방법을 통해 오토인코더 모델의 데이터의 중요 특성을 파악한다[2]. 이후 중요 특성을 이용하여 연산 복잡도가 작은 Isolation Forest 모델 기반으로 네트워크 이상탐지를 진행하는 방식이다. 또한 제안하는 방법의 유효성을 입증하기 위하여 기존 오토인코더 기반 이상탐지 방법과의 성능 비교를 진행하였다.

Ⅱ. 오토인코더 기반 네트워크 이상탐지

오토인코더 모델은 인코더 G와 디코더 F로 구성되어 있다. 인코더 G 는 입력된 데이터를 저차원으로 축소하는 역할을 하며, 원본 데이터를 압축하여 표현할 수 있도록 데이터를 축소한다. 디코더 F는 인코더에서 축소된 데이터를 다시 입력 차원으로 복원하여 출력하는 역할을 한다. 입력 데이터 $\boldsymbol{X}$ 에 대하여 오토인코더의 수식은 다음과 같다.

$$\widehat{\boldsymbol{X}} = F \circ G(\boldsymbol{X}) \tag{1}$$

수식 (1)에서 $\widehat{\boldsymbol{X}}$는 입력 데이터 $\boldsymbol{X}$에 대한 인코더와 디코더의 연산 과정을 진행한 복원 데이터를 의미한다.

오토인코더 기반 이상탐지는 정상 데이터를 사용하여 입력 데이터와 복원 데이터 사이의 복원 오차를 최소화하는 방향으로 학습을 진행한다. 따라서 인코더 G 는 정상 데이터를 가장 잘 압축하여 표현하도록 설정되며, 디코더 F 는 압축하여 표현된 데이터를 복원 오차가 최소화되게 출력하도록 설정된다. 학습에서 사용된 정상 데이터와 패턴이 상이한 비정상 데이터가 입력되었을 경우, 복원 오차가 크게 출력될 것이다. 이를 이용하여 오토인코더 기반 이상탐지는 복원 오차가 특정 임계값을 초과하는 데이터일 경우, 비정상 데이터로 판단한다.

Ⅲ. 데이터의 중요 특성 기반 네트워크 이상 탐지

본 장에서는 제안하는 데이터의 중요 특성 기반 네트워크 이상탐지 방법에 대하여 서술한다. 먼저 데이터의 특성 중요도를 파악하기 위해 복원 오차 기반의 특성 중요도 도출 방법을 서술한다. 그 뒤에 중요도가 높은 일부 특성만을 사용한 중요 특성 기반 이상탐지 방법에 대하여 서술한다.

3.1 데이터 특성 중요도 파악 방법

중요한 데이터 특성이란, 딥러닝 연산 결과에 영향을 많이 미치는 특성을 의미한다. 따라서 본 논문에서는 데이터의 특성 중요도를 파악하기 위하여 각각의 특성을 다른 값으로 교체하고, 교체 이후의 복원오차가 이전에 비해 얼마나 달라졌는지를 바탕으로 특성 중요도를 파악한다. 특성 교체 이전의 입력 데이터 셋 $\boldsymbol{X} \in \mathbb{R}^{N\times M}$는 M개의 특성을 가지는 N개의 데이터 샘플 x_n으로 이루어져 있으며 수식 (2)와 같다.

$$\boldsymbol{X} = [x_1, x_2, \cdots, x_N]^T,\ x_n \in \mathbb{R}^{M\times 1} \tag{2}$$

입력 데이터 셋 $\boldsymbol{X}$는 정상 데이터 셋 $\boldsymbol{X}_{nor}$과 비정상 데이터 셋 $\boldsymbol{X}_{abnor}$으로 구성되어 있으며 각 데이터 셋에 대한 특성 교체 이전의 복원 오차 값 $E_{1,nor}, E_{1,abnor}$은 수식 (3)과 같이 표현할 수 있다.

$$\begin{aligned} E_{nor} &= \mathbb{E}(|\boldsymbol{X}_{nor} - \widehat{\boldsymbol{X}}_{nor}|) \\ E_{abnor} &= \mathbb{E}(|\boldsymbol{X}_{abnor} - \widehat{\boldsymbol{X}}_{abnor}|) \end{aligned} \tag{3}$$

데이터 셋 $\boldsymbol{X}$의 특성 교체는 그림 1 과 같이 m번째 특성에 대한 교체를 진행 할 때, 다른 특성의 값은 입력 데이터 셋 $\boldsymbol{X}$의 값을 유지한다. 이와 같은 방식을 통해 특성 교체가 이루어진 데이터는 $\boldsymbol{X}_m^*$으로 정의한다. $\boldsymbol{X}_m^*$의 정상 데이터 셋에 대한 복원 오차 $E_{m,nor}^*$과 비정상 데이터 셋에 대한 복원오차 $E_{m,abnor}^*$은 수식

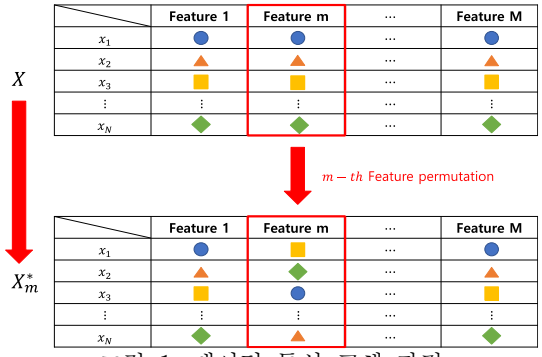

그림 1. 데이터 특성 교체 과정

(4)와 같이 표현할 수 있다.

$$E^*_{m,nor} = \mathbb{E}(|\boldsymbol{X}^*_{m,nor} - \hat{\boldsymbol{X}}^*_{m,nor}|)$$
$$E^*_{m,abnor} = \mathbb{E}(|\boldsymbol{X}^*_{m,abnor} - \hat{\boldsymbol{X}}^*_{m,abnor}|) \quad (4)$$

데이터의 특성 중요도를 파악하기 위하여 특성 교체 이전의 데이터 셋 $\boldsymbol{X}$ 의 복원 오차 값과 특성 교체 이후의 데이터 셋 $\boldsymbol{X}^*_m$ 의 복원 오차 값의 비교를 진행한다. 이 때 정상 데이터의 경우, 복원 오차 값이 크게 증가할 수록 중요한 특성이라고 판단한다. 정상 데이터가 가지고 있는 특성 값을 교체하였을 때, 복원 오차가 크게 증가한다는 것은 해당 특성이 복원 오차 값에 큰 비중을 차지하고 있기 때문이다. 반면에, 데이터 특성 교체 이후, 비정상 데이터의 복원 오차 감소는 데이터의 이상 정도의 감소와 탐지의 어려움 증가를 의미한다. 따라서 교체 이후의 복원 오차가 크게 감소할 수록 대상의 특성이 이상 탐지에 중요하게 사용된다고 판단한다. 이를 바탕으로 데이터의 m 번째 특성 중요도 D_m을 수식 (5)와 같이 정의한다.

$$d_{m,nor} = E^*_{m,nor} - E_{nor}$$
$$d_{m,abnor} = E_{abnor} - E^*_{m,abnor}$$
$$D_m = d_{m,nor} + d_{m,abnor} \quad (5)$$

$d_{m,nor}$ 은 m 번째 특성에 대한 교체 이후, 정상 데이터의 복원 오차 변화량을 의미한다. 교체 이후의 복원 오차인 $E^*_{m,nor}$ 이 교체 이전의 복원 오차인 E_{nor} 보다 커질 수록 $d_{m,nor}$ 은 큰 값을 가진다. $d_{m,abnor}$ 은 m 번째 특성에 대한 교체 이후, 비정상 데이터의 복원 오차 변화량을 의미한다. 교체 이전의 복원 오차인 E_{abnor} 이 교체 이후의 복원 오차인 $E^*_{m,abnor}$ 보다 클수록 $d_{m,abnor}$ 은 큰 값을 가진다. 최종 특성 중요도 지표 D_m 은 $d_{m,nor}$ 과 $d_{m,abnor}$ 의 합으로 표현한다.

3.2 중요 특성 기반 저복잡도 이상탐지

본 논문에서는 특성 중요도 지표 D_m 이 높은 일부 상위 특성만을 사용하는 Isolation Forest 기반의 이상탐지 방법을 제안한다. Isolation Forest 는 의사결정 나무(Decision Tree)를 응용한 머신러닝 기법 중 하나로, 정상 데이터의 분리를 위해서는 의사 결정 나무의 하단부 까지 내려가야하는 반면에, 비정상 데이터의 분리는 상단부에서 가능하다는 것을 이용한 방법이다[3].

제안하는 방법은 데이터의 중요한 특성 일부만을 사용하는 동시에 Isolation Forest 방식을 사용하여 저복잡도의 연산으로 이상탐지가 가능할 것으로 기대한다.

표 1. 성능 비교 실험 결과

		F1-score	Accuracy
Auto-encoder	mean	0.7368	0.7070
	std	0.0278	0.0263
제안 방법	mean	0.6883	0.5423
	std	0.0196	0.0454
달성률		93%	77%

Ⅳ. 성능 비교 실험

기존 오토인코더 기반 네트워크 이상탐지 성능과, 제안하는 중요 특성 Isolation Forest 기반 네트워크 이상탐지 성능의 비교를 위한 실험을 진행하였다. 성능 비교 실험을 위해 대표적인 네트워크 이상탐지 데이터인 CSE-CIC-IDS-2018 데이터를 사용하였다[4].

성능 비교를 위한 성능 평가 지표로는 Accuracy, F1-score 를 사용하였다. 표 1 은 제안 방법과 오토인코더 기반 방법의 성능을 나타낸 결과이다. 표를 통해 제안 방법의 오토인코더 성능 대비 달성률이 F1-score 는 93%, Accuracy 는 77% 인것을 확인할 수 있다. 이는 오토인코더 대비 약 $\frac{1}{10}$ 분량의 데이터만 사용하면서, 연산 복잡도가 낮은 Isolation forest 기반의 이상탐지 방식을 사용한 결과이므로 제안 방식이 저사양 기기에서도 활용 가능성이 유효함을 확인할 수 있다.

Ⅴ. 결론

본 논문에서는 데이터의 중요 특성을 사용한 Isolation Forest 기반의 저복잡도 네트워크 이상탐지를 제안하였다. 기존 오토인코더 기반 방식과의 성능 비교 실험 결과, 사용하는 데이터의 양과 연산량이 작음에도 불구하고 기존 방식 대비 93%의 F1-score, 77%의 accuracy 를 달성하였다. 이를 통해 데이터의 중요 특성을 사용한 저복잡도 이상탐지의 가능성에 대하여 제시하였을 뿐만 아니라, 데이터의 중요 특성이 오토인코더 모델의 작동 방식을 이해하는데 도움이 될 것이라고 기대할 수 있다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(No. 2021-0-00739)과 한국연구재단의 (No. NRF-2020R1F1A1069182) 지원을 받아 수행된 연구임.

참 고 문 헌

[1] 김미르, 계효선, 권민혜, “오토인코더 기반 네트워크 침입탐지 시스템에서 계층적 탐지 역할 분담에 대한 연구,” 한국통신학회 동계종합학술발표회, Feb. 2022.

[2] L. Breiman, “Random Forests,” Machine Learning, vol. 45, pp. 5-32, Oct. 2001.

[3] X. Chun-Hui, et al., "Anomaly Detection in Network Management System Based on Isolation Forest," 2018 ICNISC, Apr. 2018.

[4] https://www.unb.ca/cic/datasets/ids-2018.html