논문 21-46-06-03 The Journal of Korean Institute of Communications and Information Sciences '21-06 Vol.46 No.06
https://doi.org/10.7840/kics.2021.46.6.941

주성분 분석기반 저복잡 이상탐지 기술 연구

계 효 선◆, 권 민 혜°

PCA-Based Low-Complexity Anomaly Detection

Hyoseon Kye◆, Minhae Kwon°

요 약

사물 인터넷(Internet of Things; IoT) 기술의 발전과 함께 인터넷으로 원격 조작이 가능한 스마트 전구나 센서와 같은 초저사양 디바이스도 출시되고 있다. 이제 따라 오픈소스 플랫폼의 개발이 활발하게 이루어지고 있지만, 오픈소스는 보안의 취약점을 가지고 있다는 문제점이 존재한다. 또한, 저사양 IoT 디바이스는 CPU, 메모리 등의 낮은 하드웨어 사양으로 기존의 이상탐지 시스템을 적용하기에는 한계점이 존재하므로 저 연산, 저 복잡도, 저용량의 이상탐지 기술에 대한 필요성이 대두되고 있다. 본 논문에서는 저사양 IoT 디바이스에도 운영가능한 선형연산 기반의 저 복잡도 이상탐지 기술을 제안한다. 본 기술은 주성분 분석(Principal component Analysis; PCA)을 이용하여 정상 데이터의 주성분 벡터를 구하고, 저사양 IoT 디바이스로부터 수집된 데이터를 주성분 벡터 방향으로 선형변환 시킨 후, 새롭게 제안하는 변형된 마할라노비스 거리(Mahalanobis distance; MD)를 적용하여 이상탐지를 진행한다. 현존하는 이상탐지 기술들과의 성능을 비교한 결과 제안하는 기술이 낮은 연산 복잡도에도 불구하고 가장 높은 성능을 가지고 있음을 보였다.

Key Words : Principal component analysis (PCA), Dimensionality reduction, Anomaly detection, Mahalanobis distance

ABSTRACT

As the Internet of Things (IoT) technology has been rapidly developed, even low-end devices, such as smart light bulbs and sensors, are now able to be remotely operated via the Internet. The growth of the IoT industry leads to emerging open-source platforms, but they have a critical weakness of security vulnerabilities. Such security issue becomes much more important in low-end IoT devices. The conventional anomaly detection approaches could not be the solution in low-end devices because they have limited hardware capabilities, like low specifications of CPU and memory. Therefore, it is important to develop an anomaly detection algorithm with low computational complexity. In this paper, we propose a linear transform-based low-complexity anomaly detection solution that can be operated in low-end devices. Using principal component analysis (PCA), a principal subspace is built from normal datasets. Then, the principal components are obtained by projecting the collected IoT data onto the principal subspace. We further propose a modified Mahalanobis distance which detects anomalies from principal components. In simulation results, it is shown that the proposed solution outperforms existing methods even though it requires lower computational complexity than the others.

※ 이 논문은 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구센터지원사업(IITP-2020-2020-0-01602)과 방송통신산업기술개발사업(IIPT-2019-0-00024), 그리고 한국 연구재단(NRF-2020R1F1A1069182)의 지원을 받아 수행된 연구임.
◆ First Author : Soongsil University School of Electronic Engineering, ksh8301000@soongsil.ac.kr, 학생회원
° Corresponding Author : Soongsil University School of Electronic Engineering, minhae@ssu.ac.kr, 정회원
논문번호 : 202012-303-B-RN, Received December 7, 2020; Revised January 15, 2021; Accepted January 25, 2021

Ⅰ. 서 론

실생활에 쓰이는 대부분의 사물들을 인터넷으로 연결하는 IoT 시대가 열리면서, 인터넷을 통해 원격 조작이 가능한 스마트 전구나, 센서 등의 저사양 IoT 디바이스들이 출시되고 있다[1]. 이에 따라 저사양 IoT 디바이스를 위한 오픈소스인 ThingsBoard[2,3], IoTivity[4,5], Open Air Interface[6,7] 등의 개발도 활발하게 이루어지고 있다[8]. 이러한 오픈 소스들은 혁신적인 개발을 이루어 내지만, 보안에 매우 취약하다는 치명적인 단점이 있다. 이러한 단점을 극복할 수 있는 보안 시스템을 구축하기 위해 자동 이상탐지 기술이 연구되어 왔다[9-12].

이상탐지는 일반적이고 보편적인 데이터의 기준치에서 벗어난 데이터를 탐지해내는 방법이다. 여기서 말하는 '일반적이고 보편적인 데이터'는 정상(normal) 데이터로 간주하고, '기준치에서 벗어난 데이터'는 비정상(abnormal) 데이터로 간주한다. 방대한 양의 데이터를 활용하기 위해 대표적으로 머신러닝 방식을 사용하게 되는데, 데이터가 정상인지 비정상인지 레이블(label)을 활용하는 머신러닝의 지도학습 방식은 비정상 데이터를 취득하기 위한 시간과 비용이 많이 발생하고 정상 데이터에 비해 비정상 데이터의 비율이 현저히 적어 기존의 지도학습 알고리즘을 단순 적용하기 어렵다는 단점이 있다. 이에, 많은 연구에서 이런 단점을 보완할 수 있는 머신러닝의 비지도 학습 방식을 적용하고 있다. 머신러닝의 비지도 학습 방식은 정답이 없어도 사용할 수 있으며, 정상 데이터에 대해 학습하여 구축한 모델을 토대로 이상탐지를 진행할 수 있다. 대표적인 방식으로 PCA[9]와 오토인코더(AutoEncoder)[10]가 있다.

오토인코더는 입력된 데이터의 차원을 축소하는 인코더(encoder) 부분과 축소된 데이터를 복원시키는 디코더(decoder) 부분으로 구성된 신경망 모델이다. 복잡한 형태의 블랙박스 모델을 가지고 있어 예측에 대한 명확한 과정은 알 수 없지만, 높은 정확도를 보인다. 그러나 본 논문에서 고려하는 저사양 IoT 디바이스는 낮은 연산 능력과 적은 메모리로 구성되어 있어 오토인코더를 사용하기에는 연산 복잡도(computational complexity) 측면에서 어려움이 많다. 특히, 신경망의 깊이(레이어의 수)와 너비(한 레이어당 히든 노드의 수)가 증가함에 따라 연산 복잡도가 높아지기 때문에 저사양 IoT 디바이스에 탑재하기에는 적합하지 않다[1].

반면, PCA는 대표적인 선형 방식의 차원 축소 알고리즘이다. 데이터는 정보를 지닌 변수들을 가지고 있다. 예시로, 본 연구에서 사용하는 네트워크 트래픽 데이터는 각 패킷에서 수집한 IP, TCP, UDP 등의 정보를 담고 있다. 이처럼 수집되는 각 항목을 feature라고 명명하며, 데이터 수집 항목의 수(feature의 수)는 곧 데이터의 차원의 수를 의미한다. 차원이 높아질수록 더 정확한 정보를 얻을 수 있지만, 그만큼 분석을 하기 위해 필요한 데이터 샘플 수도 증가한다. 존재하는 데이터 샘플 수에 비해 데이터의 차원이 너무 높아지게 되면, 알고리즘 모델의 공간 복잡도(space complexity)와 시간 복잡도(time complexity)의 증가로 모델의 학습 능력이 떨어지는 차원의 저주(curse of dimensionality) 현상의 원인이 된다. 비지도 학습에서는 차원의 저주를 해결하기 위해 차원을 축소하는 기법들을 사용하여 의미 있는 정보들만 압축적으로 학습한다. 데이터의 패턴을 학습하는데 불필요한 항목은 제거하거나 다른 항목들과 결합 및 압축하여 중요한 항목들을 추출 및 조합하여 차원을 축소하게 된다. PCA 방식은 주어진 데이터의 레이블을 고려하지 않고, 흩어진 정도를 잘 나타내는 축(axis)을 분석하여 선형변환을 통해 차원을 축소하는 방식이다. 해당 축들은 원형 feature의 선형결합으로 생성되며, 이 축을 주성분 부분공간(principal subspace) 또는 선형변환(linear transformation) 벡터라고 정의한다. 부분공간은 상관관계를 가지고 있는 원형 feature에 비해 상관관계를 가지고 있지 않은 독립적인 공간으로 존재한다. 흩어진 정도가 큰 방향은 분산이 큰 방향임을 나타내고, 분산이 크다는 것은 데이터셋(dataset)의 정보를 많이 담고 있음을 의미한다. PCA를 사용하여 구한 주성분 부분공간으로 원형 데이터를 투영(projection)시켰을 때 투영 값(principal component)들의 분산이 큰 순으로 부분공간의 중요도를 결정한다. 이러한 중요도를 바탕으로 데이터의 많은 정보를 적은 수의 부분공간으로도 표현할 수 있는 최소한의 주성분 부분공간의 수를 결정한다. 이렇게 결정한 주성분 부분공간으로 원형 데이터를 투영하는 방식으로 데이터를 선형변환을 시키면, 새로운 부분공간에 데이터를 위치시킬 수 있다. 즉, 중요하다고 여겨지는 새로운 축들 만으로 고차원의 데이터를 저차원으로 축소하여 데이터를 압축적으로 분석할 수 있게 된다.

마할라노비스 거리(Mahalanobis distance; MD)는 다변량 데이터(multivariate data)에서의 상관관계(correlation)를 모두 고려하여 거리를 측정한 유사도 측정 방식이다. 즉, 각 데이터들이 전체 특징을 모두 고려했을 때, 평균으로부터 표준편차의 몇 배만큼 떨어져 있는지를 나타낸다. 하지만 기존 방식의 마할라노비스 거리 측정은 정상 데이터의 평균값과 비정상 데이터의 평균값 사이의 거리가 먼 시스템에서는 그 활용

이 적합하지 않다. 한 번에 처리되는 부분 데이터셋(subdataset)의 평균을 구하는데 정상 데이터와 비정상 데이터가 혼재되어 있으므로 그 평균값이 시스템의 정상 상태를 대표하는 값으로 결정되지 않을 수 있다는 단점이 존재한다. 즉, 정상 상태를 기준으로 하는 정확한 이상탐지가 진행되기 어렵다는 한계점을 가지고 있다. 이에 본 연구에서 제안하는 변형마할라노비스 거리(Modified Mahalanobis Distance; MMD)는 모든 데이터셋에 정상 데이터의 평균을 고려하도록 설계하였다.

본 논문에서는 저사양 IoT 디바이스에도 내재화 시킬 수 있는 선형 연산 저 복잡도 이상탐지 기술을 제안한다. 제안하는 기술은 PCA를 통해 구축한 1차원 선형변환벡터를 IoT 디바이스에 탑재한다. IoT 디바이스에서는 실시간으로 들어오는 데이터를 선형 변환시킨 후 본 논문에서 새롭게 제안하는 변형마할라노비스 거리를 기반으로 이상탐지를 수행한다. 제안된 이상탐지 기술의 성능은 현존하는 기술들과의 성능과 비교하여 낮은 연산 복잡도에도 불구하고 우월한 성능을 가지고 있음을 안드로이드 디바이스에서 수집한 네트워크 트래픽 데이터를 활용하여 확인하고자 한다.

Ⅱ. 선행 연구

네트워크 상에서의 이상탐지 연구는 활발히 진행되어왔다. 최근에 딥러닝의 발전으로 오토인코더를 사용한 이상탐지 방식이 제시되어왔다[10,13-14]. Kunang 연구팀은 차원축소와 여러 파라미터 값들을 최적화시킨 모델을 제시하며 오토인코더의 효과를 증명하였고[13], Lee 연구팀은 여러 개의 히든 레이어를 가지는 stacked 오토인코더 방식과 랜덤 포레스트(Random Forest)를 이용하여 이상탐지의 정확성과 효율성을 향상시켰다[14].

PCA를 사용한 연구 또한 활발히 진행되어왔다[9,15-20]. Lakhina 연구팀은 네트워크 데이터를 PCA를 통해 정상과 비정상의 공간을 분리하는 기법을 제시하였다[15]. Liu 연구팀은 Lakhina[15]의 알고리즘을 향상시켜 탐지 속도를 높였다.[16] 하지만 Lakhina와 Liu의 선행연구에서는 해결해야 할 2가지 문제점이 존재한다. 첫 번째는 차원 축소를 위한 적절한 주성분을 선택하는 것이고, 두 번째는 PCA를 기반으로 한 이상탐지의 성능을 높이기 위한 거리 측정 방식이다. 첫 번째 문제점에 대한 솔루션으로 적절한 주성분의 개수를 선택하기 위해 PCA를 기반으로 한 다변량 통계 네트워크 모니터링 방식과[17], 전체 데이터의 정보량의 50%에 해당하는 주성분을 선택하는 방식을 제시하였다[18]. 두 번째 문제점을 해결하기 위해 두 점 사이의 거리를 측정하는 유클리디안 거리(Euclidean Distance)[19], 민코프스키 공간(Minkowski space)에서 적용하는 민코프스키 거리(Minkowski distance) 방식[18]을 제시하였다. 또한, PCA에 공분산 행렬을 이용한 마할라노비스 거리를 적용하여 주성분 분석의 성능을 한층 높인 방식도 제시하였다[20].

앞서 언급한 선행 연구인 오토인코더 방식은 이상탐지를 위해 고성능 모델을 구축하므로 성능이 뛰어나고, PCA를 기반으로 한 선형 변환 방식 또한 높은 성능을 보인다. 하지만, 기존 연구들은 복잡도가 높아 경량 솔루션을 필요로 하는 IoT 디바이스에서는 더 이상 유효하지 않다. 본 논문에서는 IoT 디바이스에도 적용할 수 있도록, 기존 연구들보다 복잡도를 낮출 수 있는 새로운 방식을 제안한다. 제안하는 방식은 PCA를 기반으로 차원축소를 진행하며, 이상탐지 성능을 높이기 위해 마할라노비스 거리를 기반으로 변형마할라노비스 거리를 제시한다.

Ⅲ. 주성분 분석기반 저 복잡도 이상탐지 기술 제안

3.1 SVD 기반 주성분 부분공간 설계

수집된 데이터셋은 행렬 구조로 이루어져 있으며, 이는 수식 (1)과 같이 표현할 수 있다. 수집된 데이터셋은 정상 데이터와 비정상 데이터가 모두 포함되어 있다. 본 절에서는 정상 데이터를 기반으로 SVD를 이용하여 주성분 부분공간을 설계하기 위해 수집된 전체 데이터셋 (1)로부터 정상 데이터를 추출하여 (2)와 같이 별도로 표현하였다.

$$X = x_{nm}, \\ (n = 1,2, \ \cdots, N, m = 1,2, \ \cdots, M), X \in \mathbb{R}^{N \times M} \quad (1)$$

$$X_{nor} = x_{nm,nor} \ (n = 1,2, \ \cdots, N_{nor}, m = 1,2, \ \cdots, M), \\ X_{nor} \in \mathbb{R}^{N_{nor} \times M} \quad (2)$$

(1), (2)에서 X, X_{nor}의 크기는 한 번에 수집되는 데이터 항목의 수(feature의 수)인 M과 데이터 포인트의 수인 N으로 그 크기가 결정된다. $\mathbb{R}$은 실수를 의미한다. 즉, $N \times M$개의 데이터로 하나의 데이터셋이 구성된다. 정상 데이터셋인 X_{nor}은 전체 데이터 셋 X의 부분집합으로 $X_{nor} \subseteq X$를 만족한다. 데이터 항목의 수 M은 두 데이터셋이 동일하고, 정상 데이터 포인트의 수 N_{nor}는 전체 데이터 포인트 수 N보다 작거나 같다 ($N_{nor} \leq N$).

수집된 데이터는 각 항목별로 데이터 값의 범위(range)가 다양하다. 각 항목의 데이터 값의 범위를 일정하게 맞추기 위해 정규화 방식을 통해 원형 데이터의 전처리 과정을 진행한다. 정규화 과정은 수식 (3)과 같이 진행되며, 전처리 과정 후의 데이터셋은 수식 (4)와 같이 $\tilde{X}$으로 표현한다. 전처리 과정을 거친 정상 데이터셋은 $\tilde{X}_{nor} \in \mathbb{R}^{N_{nor} \times M}$로 표현하고 각 데이터 샘플은 $\tilde{x}_{nm,nor}(n=1,2,...,N_{nor}, m=1,2,...,M)$로 표현한다.

$$\tilde{x}_{nm} = \frac{x_{nm} - \mu_m}{\sigma_m}, n=1,2,...,N, m=1,2,...,M$$
$$\mu_m = \frac{\sum_{n=1}^{N_{nor}} x_{nm,nor}}{N_{nor}}, \sigma_m = \frac{\sum_{n=1}^{N_{nor}} (x_{nm,nor} - \mu_m)^2}{N_{nor}} \quad (3)$$

$$\tilde{X} = \tilde{x}_{nm}(n=1,2,...,N, m=1,2,...,M), \tilde{X} \in \mathbb{R}^{N \times M} \quad (4)$$

(3)에서는 정상 데이터 $x_{nm,nor}$의 각 항목별 평균 μ_m와 표준편차 σ_m를 이용하여 정규화 과정을 진행한다.

다음으로, 정상 데이터의 특성을 잘 나타내는 선형 변환벡터 V를 Singular Value Decomposition (SVD)를 통해 구한다. SVD는 데이터의 주성분을 분석하기 위해서 스케일링 된 정상 데이터 행렬을 수식 (5)와 같은 세 개의 행렬 곱으로 분해하는 방식이다.

$$\tilde{X}_{nor} \cong U \Sigma V^T \quad (5)$$

여기서 U는 $N_{nor} \times N_{nor}$ 크기의 좌특이벡터(left singular vector)로 직교행렬(orthogonal matrix)이고, Σ 는 $N_{nor} \times P$ 크기의 특이 값(singular value)을 가지는 대각행렬(diagonal matrix)이다. V는 $M \times P$ 크기의 우특이벡터(right singular vector)로 직교행렬(orthogonal matrix)이다. 좌특이벡터 U와 우특이벡터 V는 그 에너지 값이 항상 1이다. 즉, $UU^T = I$, $VV^T = I$를 항상 만족한다. P는 M차원의 데이터에서 축소하고자 하는 데이터 차원의 개수이다. 만약, $P = M$인 경우에는 데이터의 손실 없이 $\tilde{X}_{nor} = U \Sigma V^T$을 만족한다. $P > M$인 경우는 고려하지 않는다.

본 논문에서는 데이터의 차원을 축소하여 연산의 복잡도를 최소화시키기 위해 $P \leq M$인 경우를 고려한다. 이 경우에는 $U \Sigma V^T$로 $\tilde{X}_{nor}$을 근사적 표현을 할 수 있게 된다. 즉, $\tilde{X}_{nor} \approx U \Sigma V^T$이다. $P = M$인 경우의 우특이벡터 V를 $V = [v_1, v_2, \cdots, v_M]$와 같이 M개의 벡터들로 표현한다면, 데이터의 차원을 M에서 p로 축소시킨 $P = p < M$인 경우, $M \times p$ 사이즈의 우특이벡터 V_p는 수식 (6)과 같이 표현될 수 있다.

$$V_p = [v_1, v_2, \cdots, v_P], P = p \leq M \quad (6)$$

수식(6)에서 벡터 $v_1, v_2, \cdots, v_P$는 데이터를 근사화시키는데 가장 큰 영향을 끼치는 정도 순으로, 즉 데이터를 설명하는데 가장 중요한 벡터 순으로 정렬한다. 만약, $p = 1$이면, $V_1 = v_1$이다. 각 벡터는 데이터를 효율적으로 표현하기 위해 찾아진 새로운 주성분 부분공간의 단위방향벡터(unit vector)이다. 이는 M개의 데이터 항목별로 결정된 부분공간에 각기 다른 가중치를 주어서 새로운 부분공간으로의 회전(rotation)시키는 것으로 그 기하학적 의미를 설명할 수 있다. 새롭게 설계된 주성분 부분공간으로 투영된 데이터는 더 적은 수의 부분공간들의 조합으로도 효율적으로 원본 데이터를 표현할 수 있게 된다.

데이터의 새로운 부분공간으로의 투영과정을 그림 1을 예로 설명하고, 수식 (7)로 표현하였다.

그림 1은 $M = 2$인 2차원 데이터를 그래프 상에 나타내었다. 수식(5)의 SVD 방식를 사용하여 중요도가 가장 큰 한 개의 주성분 부분공간을 나타내었다. 이 주성분 부분공간으로 원본 데이터를 투영시킨 투영 값들을 통해 데이터의 차원이 1차원으로 축소됨을 알 수 있다.

$$\hat{X} = \tilde{X} \bullet V_p, 0 < p \leq M, p \in \mathbb{N} \quad (7)$$

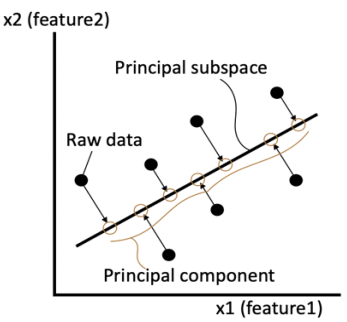

그림 1. 주성분 축과 데이터 투영
Fig. 1. Principal subspace and data projection

$\mathbb{N}$은 자연수를 의미한다. 이러한 과정은 기존의 데이터 $\tilde{X}$가 선형변환 벡터 V_p를 통해서 데이터 $\hat{X}$로 선형변환 되었다고 해석될 수 있다. 수식(7)에서 $\hat{X}$는 $\mathbb{R}^{N\times p}$로 데이터셋의 차원이 $N\times M$에서 $N\times p$로 축소된다.

본 연구에서는 IoT 디바이스에서의 연산 복잡도를 최소화하기 위해서, 데이터의 흩어진 정도를 가장 잘 나타내는 첫 번째 주성분 부분공간 만을 사용하여 차원을 축소하고자 한다. 즉, $p=1$로 설정하고자 한다. 이를 위해 수식(5)에서 찾은 V의 첫 번째 열벡터 $v_1\left(v_1\in\mathbb{R}^{M\times 1}\right)$만으로 V_1을 구성한 후, 해당 선형변환 벡터로 $\tilde{X}$를 변환하여 $\hat{X}$를 구성한다. 이 과정은 수식(8)과 같다.

$$\hat{X}=\tilde{X}\bullet\ V_1 \tag{8}$$

수식(8)에서 $\hat{X}$는 $N\times 1$ 사이즈로 데이터의 차원이 축소된다.

시스템 설계 시, 수집된 정상 데이터를 바탕으로 V_1을 디자인 해놓고, IoT 디바이스에서 이상탐지 수행 시에는 단순 선형연산인 수식(8)만을 수행한 후, 다음 절에서 제안하는 변형마할라노비스 거리를 기반으로 이상 여부를 탐지한다.

3.2 변형마할라노비스 거리 기반 이상탐지 기술

저사양 IoT 디바이스에 네트워크 트래픽이 발생되면 수식(3)의 정규화 과정을 거친 후, 수식(8)에서 제안한 선형 변환을 수행하여 $\hat{X}$를 구한다. 여기서 $\hat{X}$는 정상 데이터에만 한정되어 있지 않고 비정상 데이터 또한 포함될 수 있다는 것이 앞 절과의 차이점이다. 선형변환된 $\hat{X}$를 바탕으로 본 논문에서 제안하는 변형마할라노비스 거리를 측정($f_{MMD}(\hat{X})$)하여 임계점 δ보다 거리가 큰지를 기준으로 이상여부를 판별하는데, 이는 수식(9)와 같다.

$$f_{MMD}(\hat{X})=\sqrt{(\hat{x}_i-\mu)\,COV^{-1}(\hat{X})(\hat{x}_i-\mu)^T},\quad i=1,2,...,N$$
$$COV^{-1}(\hat{X})=\frac{(\hat{X}-\mu)^T(\hat{X}-\mu)}{N},\ \mu=E(\hat{X}_{nor}) \tag{9}$$

여기서 μ는 시스템 설계 시 정해 놓은 상숫값으로, 정상 데이터 $\hat{X}_{nor}$의 평균값으로 설정한다. 즉, $\mu=E(\hat{X}_{nor})$이다. 제안하는 $f_{MMD}(\hat{X})$는 입력되는 데이터 $\hat{X}$가 정상 데이터의 평균값 μ로부터 몇 배의 표준편차만큼 떨어져 있는지를 측정한다. 즉, 데이터 정상 범주로부터 얼마나 떨어져 있는지를 정량적으로 측정할 수 있다.

마할라노비스 거리를 측정하여 임계점을 설정하기 위해, 선행 연구에서는 경험적으로 임계점을 결정하는 방식을 사용하였다[18,21-25]. 경험적인 규칙(Empirical Rule)을 사용하여 표준편차의 2배수[18] 혹은 3배수[21]를 임계점으로 설정 한다. 혹은, 정상과 비정상의 마할라노비스 거리를 측정하여, 정상 데이터의 분포에 따르지 않는 경우를 비정상으로 정의하여 이상탐지를 진행한다[22-25]. 본 논문에서 제안한 이상탐지 기술을 적용하는 선형변환된 네트워크 데이터는 정규분포를 따르지 않으므로, [22-25]에서 사용한 방식과 같이 정상 데이터의 분포에 따르지 않는 데이터를 비정상 데이터로 간주하였다. 즉, 선형변환된 정상 데이터의 마할라노비스 거리 분포를 기준으로 경험적으로 임계점을 설정하였다.

기존의 마할라노비스 거리 측정 방식은 정상 데이터에 한정되지 않고 비정상 데이터와 혼재된 데이터셋의 평균값을 μ로 사용하므로 수식(9)에서 $\mu=E(\hat{X})$이 적용되었다. 기존의 방식에서는 비정상 데이터의 평균값과 정상 데이터의 평균값의 거리가 먼 경우 대표성을 띄지 못하는 임의의 μ값이 선택될 수 있다는 한계점이 있었다. 또한, 처리하고자 하는 데이터셋 내에 포함되는 비정상 데이터의 비율에 따라 μ값이 변화한다는 한계점이 있었다. 이는 그림 2에서 설명하고 있다. 반면, 본 논문에서 제안하는 변형마할라노비스 거리 측정 방식은 이러한 한계점을 극복할 수 있도록 정상 데이터의 평균값으로 μ를 고정하였다. 이는 그림 3에서 설명한다.

그림 2와 그림 3에서, 파란색 타원은 전체 데이터 중 정상 데이터를 의미하고, 빨간색 타원은 비정상 데이터를 의미한다. 검은색은 각 그림에 적용된 평균 값 μ를 나타내며, 녹색은 정상 데이터와 가까운 위치에 있어 탐지가 어려운 비정상 데이터를 나타낸다. 하늘색은 같은 표준편차 거리 값을 가지고 있는 위치를 나타내며, 그림 3은 특히 정상 데이터의 임계점 위치를 나타냈다. 비정상 데이터인지를 판별하는 것은, 임계점 δ를 바탕으로, $f_{MMD}(\hat{X})>\delta$일 경우를 비정상으로 판단한다. 여기서 δ는 사전에 정의된 상숫값으로, 시스템 설계 시 정상 데이터의 흩어짐 정도를 바탕으로 결정한다. 그림 2는 전체 데이터셋의 평균값을 μ로 사용하므

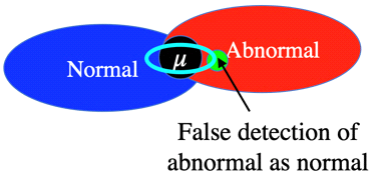

그림 2. 기존의 마할라노비스 거리
Fig. 2. Conventional Mahalanobis Distance

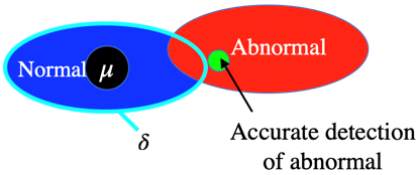

그림 3. 변형마할라노비스 거리
Fig. 3. Modified Mahalanobis Distance

로 정상 데이터와 비정상 데이터가 평균으로부터 같은 표준편차 거리에 존재할 가능성이 높다. 따라서 기존의 마할라노비스 거리 방식은 임의의 값 μ로 인해 명확한 기준으로 이상탐지를 진행할 수 없으므로 비정상 데이터(녹색)를 정상으로 탐지하는 오류를 일으킬 가능성이 크다는 것을 알 수 있다. 그림 3은 정상 데이터의 평균값 μ를 기준으로 정상 데이터의 임계점 δ를 정의하였다. δ를 벗어난 데이터는 비정상으로 탐지되므로 정상 데이터에 근접해있는 비정상 데이터(녹색)도 비정상으로 탐지하여 정확한 이상탐지를 진행할 수 있다.

Ⅳ. 제안된 이상탐지 기술의 성능 평가

본 논문에서 제안하는 저사양 IoT 디바이스를 위한 이상탐지 기술의 성능을 평가하기 위해서 대표적인 오픈소스인 안드로이드를 탑재한 모바일 디바이스에서 수집된 네트워크 데이터를 바탕으로 이상탐지를 수행한다. 제안된 이상탐지 기술은 PCA 기반으로 단 하나의 주성분 부분공간을 설계하여 데이터를 선형 변환시켜 차원 축소를 진행하고, 선형변환된 데이터의 변형마할라노비스 거리를 측정하여 임계치보다 높으면 비정상임을 탐지하는 방식이다. 본 논문에서는 7가지의 성능지표를 사용하여 4가지의 다른 이상탐지 기법과 성능을 비교하여 제안된 이상탐지 기술 성능의 우수성을 평가하고자 한다.

모바일 디바이스에서 수집된 네트워크 데이터를 사용하여 평가한 제안된 이상탐지 기술 성능의 우수성을 바탕으로 실제 IoT 디바이스에서 수집된 네트워크 데이터에 제안된 이상탐지 기술의 성능을 검증한다. 본 제안 기술은 IoT 디바이스에 적용하는 것을 목적으로 하므로 7가지의 성능 지표와 3가지의 다른 이상탐지 기법과 성능을 비교하여 제안된 이상탐지 기술 성능의 우수성을 평가하고자 한다.

4.1 안드로이드 기기에서 수집된 네트워크 데이터 분석

본 연구에서는 오픈 소스 특성으로 인해 보안에 취약한 안드로이드 디바이스에서 수집된 네트워크 데이터를 바탕으로 이상탐지를 진행한다. 본 연구에서 사용한 데이터는 안드로이드 디바이스에서 Droid Collector[26,27]를 통해 수집된 네트워크 데이터이다. 해당 데이터는 kaggle에 공개되어있다[28]. DroidCollector는 안드로이드 트래픽 수집을 위한 모바일 애플리케이션으로 패킷 캡처(PCAP) 파일을 제공해주는 고성능 프레임 워크이다. PCAP 파일은 네트워크 트래픽을 캡처하기 위한 API로 구성되어 있어, 네트워크 트래픽을 분석하기에 많이 사용되는 라이브러리이다. 이 프레임 워크를 통해 약 140시간 동안 5,560개의 악성 앱에서 생성된 330MB 트래픽데이터와 6,000개의 무해한 앱에서 생성된 808MB의 트래픽 데이터를 수집하였고, 본 연구에서 사용한 데이터 샘플의 수는 7845개이며, 각 데이터 샘플은 13개의 수집 항목으로 구성되어 있다. 즉, 사용한 데이터셋은 7845×13개의 데이터로 구성되어 있다. 또한, 데이터 샘플은 4704개의 정상 데이터와 3141개의 비정상 데이터로 구성되어 있다.

데이터 샘플의 각 수집 항목들에 대한 정보는 표 1과 같다. 인터넷 프로토콜(Internet Protocol; IP)은 송/수신 호스트 간에 패킷 교환 네트워크에서 정보를 주고받기 위해 사용하는 통신 프로토콜이다. 패킷을 전송할 때 패킷의 전송을 보장하고, 전송한 패킷의 순서를 보장하기 위해 IP의 상위 프로토콜인 TCP 프로토콜을 사용한다. HTTP, FTP, SMTP 등의 대부분의 네트워크 애플리케이션 프로토콜들은 신뢰성과 전송 순서를 보장하는 TCP를 기반으로 한다. 데이터 전송 시에 DNS, SNMP, RIP등의 신뢰성과 전송 순서를 보장하지 않아도 되는 네트워크 어플리케이션에는 UDP 프로토콜을 사용한다. 또한, DNS는 클라이언트가 도메인의 IP주소를 요청할 때, 도메인을 IP주소로 변환하는 역할을 한다. 표 1의 수집 항목들은 데이터의 송/수신을 위한 네트워크 프로토콜을 통해 수집되는 항목들이다.

본 논문에서 제안하는 방식은 비지도 학습 방식이기에, 데이터의 종류를 의미하는 'type' 항목은 제외한 총 12개의 수집 항목들이 학습 및 이상탐지 수행 시에 사용되었다. 'type' 항목은 시스템 설계 및 성능 평가

표 1. 데이터 수집 항목의 정보
Table 1. Description of data features

	Feature	Description	Value [Bytes]
1	tcp_packets	The number of TCP packets sent and received during comnication	2B
2	dist_port_tcp	The number of packets with ports different from those exposed to TCP	1.5B
3	external_ips	The number of external IP addresses which communicate with the application	1B
4	vulume_bytes	Bytes transmitted from the application to the external site	5MB
5	tcp_urg_packet	The number of Urgent packets over TCP	0.2B
6	udp_packets	The number of UDP packets transmitted in communication	1B
7	source_app_packets	The number of packets sent from the application to the remote server	2B
8	remote_app_packets	The number of packets received from remote server to the application	2B
9	source_app_bytes	Network traffic bytes between the application and the remote server	66MB
10	remote_app_bytes	Data volume from the remote server to the application	5MB
11	source_app_packets_1	The number of packets sent from the application to the remote server	2B
12	dns_query_times	DNS query call times	1B
13	Type	Data type (normal/abnormal)	9B

용도로만 사용하였다. 3.1절에서 시스템 설계 시에 정상 데이터셋을 구성하기 위해 전체 데이터셋에서 'type' 항목을 사용하여 정상 데이터를 추출하였다. 또한 본 논문에서 제안한 시스템의 성능을 평가하기 위해 4.2절에 기술한 오차 행렬의 'True condition'에 'type' 항목을 사용하였다.

해당 데이터의 특성은 다음과 같은 두 가지 시각화(visualization) 과정을 통해 확인하였다.

• 데이터 수집 항목(feature) 기반 시각화

PCA 등의 추가 연산 없이 규칙기반(rule-based)방식으로 이상탐지가 가능한지 여부를 확인하기 위하여, 데이터 항목의 특정 값을 기준으로 규칙(rule)을 만들어 그림 4, 그림 5와 같이 데이터를 시각화하였다. 그림 4와 그림 5의 x축은 표 1의 1번 수집 항목인 'tcp_packets'를 의미하고, 그림 4의 y축은 표 1의 3번 수집항목인 'external_ips', 그림 5의 y축은 표 1의 8번 수집 항목인 'remote_app_packets'를 의미한다. 정상 데이터는 파랑색, 비정상 데이터는 빨간색으로 표시하였으며, 두 그림 모두에서 공통적으로 비정상 데이터가 정상 데이터와 겹쳐서 표시되기에 구분이 어려움을 확인하였다.

• 비선형 변환 기반 시각화: t-SNE

대표적인 고차원데이터의 시각화 방식인 t-SNE[29]를 바탕으로 그림 6에 12차원 데이터를 시각화하였다. 빨간색은 정상, 파란색은 비정상 데이터를 나타낸다. t-SNE는 비선형 방식으로 각 데이터의 분포에서 데이터 간의 거리를 측정한다. 이 데이터 간의 거리를 친밀도(similarity)라고 하고, 이 친밀도가 가까운 데이터끼리 묶어 시각화시킨다. 그림 6을 통해 비선형 방식인 t-SNE로 데이터를 시각화한 결과, 정상 데이터와 비정

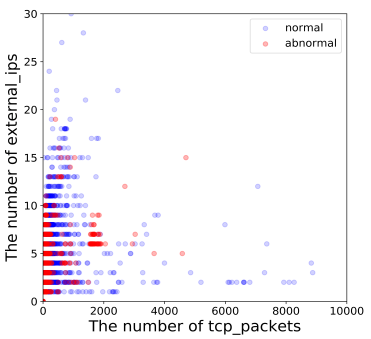

그림 4. 데이터 수집항목 1
Fig. 4. Data feature 1

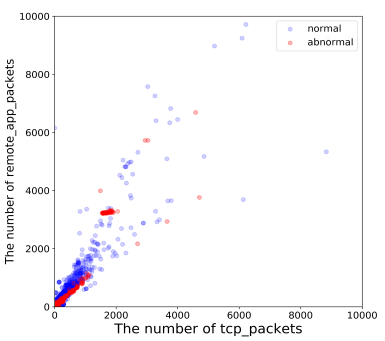

그림 5. 데이터 수집항목 2
Fig. 5. Data feature 2

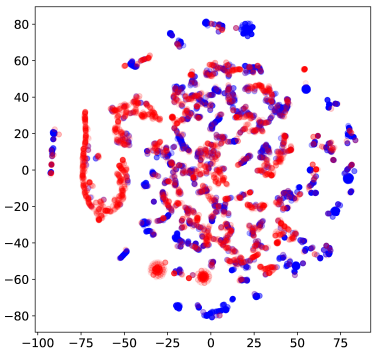

그림 6. t-SNE (파란색: 정상, 빨간색: 비정상)
Fig. 6. t-SNE (blue: normal, red: abnormal)

상 데이터가 쉽게 구분되지 않고 많은 부분 겹쳐서 표현되는 것을 확인할 수 있다.

이렇게 본 논문에서 다루는 데이터는 고차원의 공간에서 정상과 비정상 데이터가 혼재되어 분포되기에 그 구분이 매우 어렵다는 특징이 있다. 다음 장에서는 본 논문에서 제안하는 이상탐지 기술을 적용하였을 때의 성능을 확인한다.

4.2 비교 성능 평가 분석

제안하는 이상탐지 기술방식의 성능을 안드로이드 기기에서 수집된 네트워크 트래픽 데이터를 바탕으로 이상탐지를 수행하여 평가하고 현존하는 4가지의 이상탐지 기술방식의 성능과 비교 한다. 성능 평가는 이상탐지 기술 평가에 널리 사용되는 다음의 7가지 평가 지표를 사용한다.

먼저 평가 지표에 활용되는 오차 행렬(confusion matrix)을 표 2와 같이 정의한다. Positive는 예측한 값이 비정상일 경우를 나타내고, Negative는 예측한 값이 정상일 경우를 나타낸다. True는 예측이 맞는 경우를 나타내고, False는 예측이 틀린 경우를 나타낸다.

표 2. 오차 행렬
Table 2. Confusion matrix

		True condition	
		Positive	Negative
Predicted condition	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

False Positive (FP): 정상 데이터가 비정상으로 예측된 데이터의 개수
False Negative (FN): 비정상 데이터가 정상으로 예측된 데이터의 개수
True Positive (TP): 비정상 데이터가 비정상으로 예측된 데이터의 개수
True Negative (TN): 정상 데이터가 정상으로 예측된 데이터의 개수

• 정밀도(precision)

정밀도는 예측 결과가 비정상인 것 중에서 실제 정답이 비정상인 비율이다. 이 값이 1에 가까울수록 오경보(false alarm)의 비율이 적어져서 시스템의 효율성이 높아진다.

$$Precision = \frac{TP}{TP+FP}$$

• Recall

Recall은 실제 정답이 비정상인 것 중 예측 결과 또한 비정상인 비율이다. 이 값이 1에 가까울수록 비정상인데 정상이라고 판단하여 시스템에 악영향을 주는 케이스가 줄어들어서 더욱 안정적인 시스템 운영이 가능하게 된다.

$$Recall = \frac{TP}{TP+FN}$$

• F1 Score

F1 Score는 정밀도와 Recall의 조화평균이다. 정밀도가 높아지면 오경보의 경우가 낮아져 시스템의 효율이 향상될 수 있지만, 동시에 비정상으로 판단하는 횟수 자체가 줄어들어 Recall이 낮아져 시스템이 더 큰 위험에 노출될 수 있다. 따라서 정밀도와 Recall사이에 적절한 밸런스를 맞추는 이상탐지 시스템을 구축하는 것이 중요한데, 이를 평가하는 것이 F1 Score이다. 이상탐지 시스템의 성능을 평가하는데 가장 많이 활용되고 있는 성능 지표 중 하나이다.

$$F1\ \ Score = \frac{2(Precision \times Recall)}{Precision + Recall}$$

• 정확도(accuracy)

정확도는 전체 데이터에서 모델이 예측한 결과와 실제 정답이 같은 비율이다. 이 값이 1에 가까울수록 이상탐지 기술의 예측이 정확하다는 것을 의미한다.

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN}$$

• Fall out

Fall out은 실제 정상인 데이터 중에서 예측 결과가 비정상인 비율이다. 이 값이 커질수록 오경보의 비율이 늘어나는 것이기에 시스템의 효율성은 떨어진다.

$$Fall\ \ out = \frac{FP}{TN+FP}$$

• Specificity

Specificity는 실제 정상인 데이터 중 예측 결과가 정상인 데이터의 비율이다.

$$Specificity = \frac{TN}{TN+FP}$$

• Matthews Correlation Coefficient (MCC)

MCC는 -1에서 1사이의 값을 가지며, 1에 가까울수록 예측이 정확하다는 의미이고, -1에 가까우면 예측이 정확히 반대로 이루어졌다는 의미이며, 0에 가까울수록 랜덤 예측을 의미한다. 정상 데이터 수와 비정상 데이터 수가 큰 차이를 보일 때에는 정확도 측정 방식이 적절치 못하다는 한계점을 극복하기 위해 MCC가 제안되었다[30]. 일반적인 시스템에서는 정상 데이터의 수가 비정상 데이터의 수보다 훨씬 많기 때문에 정확도 뿐만 아니라 MCC값을 확인하는 것이 중요하다.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)\times(TP+FN)\times(TN+FP)\times(TN+FN)}}$$

위에서 설명한 7가지의 평가 지표를 사용하여 본 연구에서 제안하는 이상탐지 기술의 성능을 평가하고 그 결과를 그림 7과 표 4에 명시하였다. 안드로이드 디바이스에서 수집된 네트워크 트래픽 수식(3)을 통해 데이터 전처리 과정을 거치고, 수식(5)의 SVD 과정을 통해 선형변환 벡터 V_1을 구한 후, 수식(8)의 방식으로 선형변환 된 $\hat{X}$를 구하였다. $\hat{X}$의 차원은 $\mathbb{R}^{N\times 1}$로 결과적으로 12차원의 원형 데이터를 1차원으로 축소시켰다. $\hat{X}$에 제안된 수식(9)의 변형마할라노비스 거리를 이용하여 이상여부를 탐지하였고 이 때 임계값 $\delta = 0.25$를 사용하였다. 그림 7은 변형마할라노비스 거리에 따른 정상 데이터와 비정상 데이터의 수(frequency)를 나타내었다. 변형마할라노비스 거리가 0.25일때를 기준으로 정상 데이터의 수는 급격히 줄어들고 비정상 데이터의 수는 늘어남을 확인할 수 있다.

제안하는 이상탐지 기술과 그 성능을 비교하기 위하여 다음의 네 가지 이상탐지 기술을 통해 그 성능을 확인하였다.

• 비교 방식 1: PCA를 이용하여 1차원으로 축소한 후 기존의 마할라노비스 거리를 적용하여 이상탐지 수행 (PCA, 1dimension, MD)

본 논문에서 제안하는 변형마할라노비스 거리가 기존 마할라노비스 거리의 한계를 보완한다는 것을 확인하기 위해, 제안된 방식과 모든 방식을 동일하게 하고 수식(9)에서 기존의 마할라노비스 거리방식을 사용하여 이상여부를 탐지하였다. 기존의 마할라노비스 거리 방식은 전체 데이터셋의 평균값을 μ로 사용하므로 수식(9)에 $\mu = E(\hat{X})$을 적용한다. 제안된 방식과 동일한 조건으로 성능 평가를 진행해보기 위해, 동일한 임계값 $\delta = 0.25$를 사용하였다. 그림 8은 기존 마할라노비스 거리에 따른 정상 데이터와 비정상 데이터의 수(frequency)를 나타내었다. 정상과 비정상 데이터 모두 거리가 0.25이하인 부분에 많이 분포하기 때문에 정상과 비정상의 구분이 어려움을 확인할 수 있다. 기존의 마할라노비스 거리방식은 데이터셋에서의 정상 및 비정상 데이터의 비율에 따라 그 평균값이 달라지기에 성능의 변화가 크다는 한계점이 있다. 이러한 한계점을 확인하기 위하여 정상 데이터와 비정상 데이터를 다양한 비율로 혼합하였을 때의 성능을 분석한 결과를 표 3에 정리하였다.

표 3은 앞에서 기술한 7가지 성능 지표를 사용하여

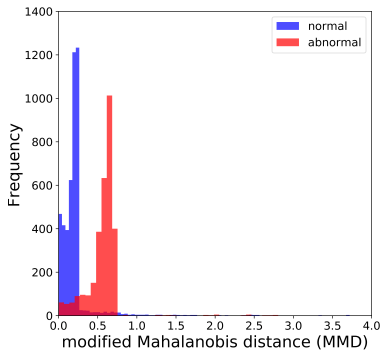

그림 7. PCA(1차원), MMD
Fig. 7. PCA (1dimension), MMD

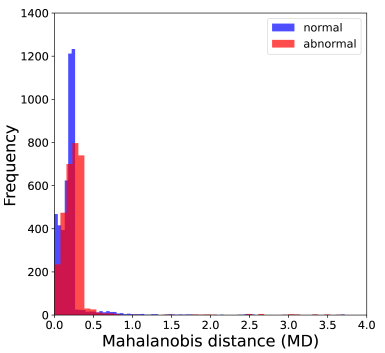

그림 8. PCA(1차원), MD
Fig. 8. PCA (1dimension), MD

표 3. 데이터 구성 비율에 따른 이상탐지 성능 평가 결과
Table 3. The performance of anomaly detection

	Precision	Recall	F1 score	Accuracy	Fall out	Specificity	MCC
normal 90% abnormal 10%	0.043	0.134	0.065	0.901	0.079	0.921	0.032
normal 70% abnormal 30%	0.276	0.532	0.364	0.893	0.085	0.915	0.332
normal 50% abnormal 50%	0.493	0.672	0.569	0.891	0.083	0.917	0.516

표 4. 안드로이드 네트워크 데이터의 이상탐지 성능 평가 결과
Table 4. The performance of anomaly detection

	Precision	Recall	F1 score	Accuracy	Fall out	Specificity	MCC
Proposed solution (PCA, 1dimension, MMD)	0.888	0.928	0.907	0.924	0.078	0.922	0.844
Comparison 1 (PCA, 1dimension, MD)	0.815	0.516	0.632	0.759	0.078	0.922	0.493
Comparison 2 (NLPCA, 1dimension, MMD)	0.866	0.923	0.894	0.912	0.095	0.905	0.820
Comparison 3 (PCA, 12dimension, MMD)	0.723	1.000	0.839	0.846	0.256	0.744	0.733
Comparison 4 (rule-based, 1dimension)	0.251	0.223	0.236	0.421	0.446	0.554	-0.228

3가지 비율로 데이터 구성이 되었을 때의 이상탐지 기술의 성능을 정량적으로 평가한 결과이다. 데이터의 혼합비율에 따라 그 성능이 상이하게 됨을 확인함으로써 기존의 마할라노비스 거리 방식의 한계점을 확인하였다.

- 비교 방식 2: NLPCA (Nonlinear PCA)를 이용하여 1차원으로 축소한 후 제안하는 변형된 마할라노비스 거리를 적용하여 이상탐지 수행 (NLPCA, 1dimension, MMD)

데이터가 비선형성을 특징으로 가질 때에 적용하는 Nonlinear PCA (NLPCA) 방식 중 오토인코더를 사용하여 비선형 변환과 차원 축소를 진행하였다. 오토인코더는 차원을 축소하는 인코더(encoder)와 다시 축소된 차원을 확장하여 원본 데이터를 복원시키는 디코더(decoder)로 구성되어 있다. 본 논문에서는 비선형함수인 ELU를 활성화 함수로 사용하는 12-1-12 차원구조의 2-layer 오토인코더 모델을 구축하여 정상 데이터를 기반으로 학습시킨 후, 인코더 출력물인 1차원 데이터에 본 논문에서 제안한 변형마할라노비스 거리를 사용하여 이상탐지를 진행하였다. 이 때, 임계값 $\delta = 0.25$를 사용하였다.

그림 9는 변형마할라노비스 거리에 따른 정상 데이터와 비정상 데이터의 수를 나타내었다. 정상 데이터가 거리가 0.25일때를 기준으로 급격히 줄어들고 비정상 데이터는 0을 기준으로 늘어나기 때문에 대체적으로 비정상과 정상 데이터의 구분이 가능하였다. 따라서 제안하는 변형마할라노비스 거리를 이용하면 선형변환 혹은 비선형 변환여부에 관계 없이 대체적으로 이상탐지가 가능함을 확인하였다. 그러나 거리가 0에서 0.25 사이에 위치한 비정상 데이터의 경우는 정상 데이터로 예측되기 때문에 표 4에 명시된 정밀도 측면에서는 그

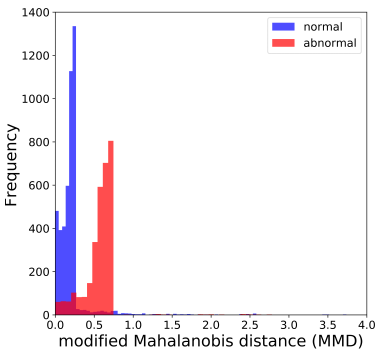

그림 9. NLPCA(1차원), MD
Fig. 9. NLPCA (1dimension), MMD

성능이 0.866로 제안된 기술의 성능인 0.888에 비해 저조하다는 한계점이 있다. 특히 정상과 비정상 데이터의 수 차이까지 고려한 MCC 값은 0.820으로 제안된 기술 0.844에 비해 훨씬 낮은 값을 가지는 것을 확인하였다.

- 비교 방식 3: PCA를 이용하여 데이터의 차원을 축소 없이 선형변환만 진행한 후 제안하는 변형마할라노비스 거리 적용하여 이상탐지 수행 (PCA, 12dimension, MMD)

본 논문에서 사용한 차원 축소 방식이 계산복잡도를 줄일 수 있는 동시에 이상여부를 판단할 수 있는 데이터 고유 정보의 양은 잘 보존되는지를 확인하기 위하여, 데이터의 차원을 축소하지 않고, PCA를 통해 구한 12개의 주성분 부분공간들을 모두 사용하여 데이터를 선형변환 시킨 뒤, 제안된 방식과 동일하게 변형마할라노비스 거리를 기반으로 이상탐지를 수행하였다. 즉, 수식(5)의 SVD를 이용하여 선형변환 벡터 V를 구한 후, 수식(7)의 방식으로 선형변환 된 $\hat{X}$를 구한다. $\hat{X}$의 차원은 $\mathbb{R}^{N\times 12}$로 차원 축소가 이루어지지 않았다는 것이 앞에서 제안하는 방법과의 차이점이다. 수식(7)을 통해 구해진 $\hat{X}$에 제안하는 변형마할라노비스 거리를 적용하여 이상여부를 탐지하였고, 이 때, 임계값 $\delta = 1.1$을 사용하였다.

그림 10은 변형마할라노비스 거리에 따른 정상 데이터와 비정상 데이터의 수(frequency)를 나타내었다. 변형마할라노비스 거리가 1.1일때를 기준으로 비정상 데이터의 수는 급격히 늘어남을 확인할 수 있다. 차원 축소가 일어나지 않기에 손실되는 정보가 없어서 비정상 데이터가 정상 데이터로 검출되는 경우는 존재하지 않다는 장점, 즉 $recall = 1$이라는 장점이 있다. 그러나, 정상 데이터가 비정상으로 검출되는 비율인 fall out이 0.256으로 제안된 기술이 0.078인 것에 비해 매우 높게 검출됨을 확인했다. 즉, 오경보 비율이 높게 검출됨을 확인했다. 즉, 오경보 비율이 높아 시스템의 효율성이 낮아지게 된다는 단점이 존재한다. 일반적으로 데이터를 구성하는 특징들을 많이 사용할수록 계산량은 높아지지만 더 좋은 성능을 보이는 경우가 많으나, 본 논문에서 사용하는 데이터는 fall out 측면에서 낮은 성능을 보여 높은 오경보율을 가져오는 것을 확인하였다.

- 비교 방식 4: 데이터의 수집 항목 중 이상탐지가 가장 잘 되는 1개만으로 규칙기반 이상탐지 수행 (rule-based, 1dimension)

이 비교방식에서는 데이터에서 이상탐지가 가장 잘 될 수 있는 1개의 수집항목만으로 이상탐지를 진행하였다. 여러 항목을 기준으로 실험한 결과, 표 1의 'tcp_packets' 항목(통신 중 송수신된 TCP 패킷의 수)이 가장 이상여부 구분을 잘 하는 것으로 확인하였다. 그림 11은 통신 중 송수신된 TCP 패킷의 수에 따른 정상 데이터와 비정상 데이터의 수(frequency)를 나타내었다. 정상과 비정상 데이터 모두가 같은 분포양상을 보이기 때문에 정상과 비정상의 구분이 어려움을 확인할 수 있다.

표 4는 앞에서 기술한 7가지 성능 지표를 사용하여 5가지 방식의 이상탐지 기술의 성능을 정량적으로 평가한 결과이다. 제안된 기술은 모든 평가 지표에서 대체적으로 가장 우수한 성능을 가지고 있음을 확인하였다. 제안된 기술이 Recall 항목에서만 두 번째로 높은 성능을 보이지만, 비교 방식 3(PCA, 12dimension, MMD)의 경우에는 fall out과 MCC 항목의 성능이 제안된 방식에 비해 현저히 낮기 때문에 종합적인 판단 아래에서는 제안된 방식이 더욱 우수하다고 결론 지을 수 있다. 또한, 기존의 마할라노비스 거리 방식으로 이

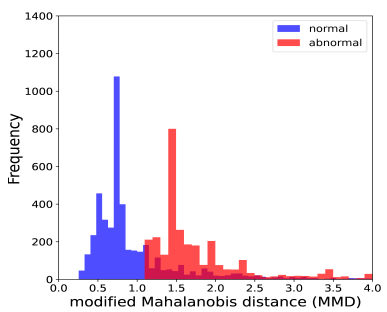

그림 10. PCA(12차원), MD
Fig. 10. PCA (12dimension), MD

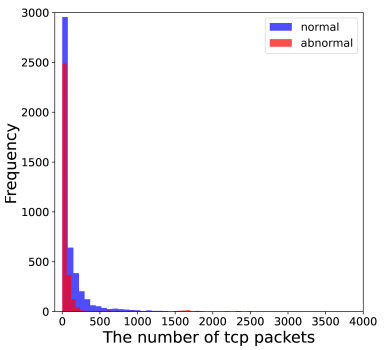

그림 11. 규칙기반 이상탐지
Fig. 11. Rule-based Anomaly detection

상탐지를 수행하는 비교 방식 1(PCA, 1dimension, MD)과 비교하여 본 논문에서 제안하는 방법의 성능 평가 결과들이 월등히 높은 것을 근거로, 변형마할라노비스 거리의 우수성을 확인하였다. 이러한 성능 향상은 거리를 측정하는 기준점을 정상 데이터의 중점으로 결정하여 측정하고자 하는 데이터가 정상 범주로부터 얼마나 떨어져 있는지를 더욱 정확하게 확인하였기 때문이다.

4.3 IoT 기기에서 수집된 네트워크 데이터 성능 평가 분석

4.1절의 안드로이드 기기 네트워크 데이터를 사용하여 4.2절에서 제안된 이상탐지 기술의 우수성을 입증하였다. 본 제안 기술이 실제 IoT 기기에 적용 가능함을 입증하기 위해 IoT 기기에서 수집된 네트워크 데이터를 바탕으로 이상탐지를 진행한다. 본 절에서 사용한 데이터는 IoT 디바이스에서의 네트워크를 기반으로 침입 탐지 시스템을 구축하기 위해 수집된 네트워크 데이터셋이다. 해당 데이터는 아두이노와 NodeMCU가 들어 있는 초음파 센서에서 네트워크를 모니터링하고 네트워크 로그를 수집하였으며 kaggle 에 공개되어있다[31]. 해당 데이터 샘플 수는 477,426개이며, 각 데이터 샘플은 14개의 수집 항목으로 구성되어 있다. 즉, 사용한 데이터셋은 477,426×14개의 데이터로 구성되어 있다. 또한, 데이터 샘플은 79,035개의 정상 데이터와 398,391개의 비정상 데이터로 구성되어 있다. 대표적인 고차원데이터의 시각화 방식인 t-SNE[29]를 바탕으로 그림 12에 14차원 데이터를 2차원으로 시각화하였다. 빨간색은 정상, 파란색은 비정상 데이터를 나타낸다. 그림 12를 통해 비선형 방식인 t-SNE로 데이터를 시각화한 결과, 정상 데이터와 비정상 데이터가 쉽게 구분되지 않고 많은 부분 겹쳐서 분포되어 있음을 확인하였다. 이에 본 제안 기술을 이용하였을 때의

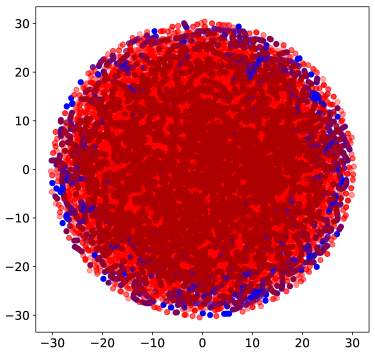

그림 12. t-SNE (파란색: 정상, 빨간색: 비정상)
Fig. 12. t-SNE (blue: normal, red: abnormal)

비교 성능을 확인한다.

표 5는 4.2절에서 사용한 7가지 성능 지표와 제안 방식, 비교 방식 1, 비교 방식 2, 비교 방식 3의 기술 성능을 정량적으로 평가한 결과이다. 제안 방식이 모든 성능 지표에서 비교 방식 1(PCA, 1dimension, MD)보다 높은 성능을 보이는 것을 근거로, 제안된 변형마할라노비스 거리의 우수성을 확인하였다. 비교 방식 2(NLPCA, 1dimension, MMD)와 비교 방식 3(PCA, 12dimension, MMD)는 Specificity에서 제안 방식 이상의 성능을 보이고 있다. 하지만, Precision과 Recall를 종합적으로 고려하는 F1 score의 값이 제안 방식보다 현저히 낮다. 이는 제안 방식이 탐지하지 못할 경우 더 치명적일 수 있는 비정상 데이터의 탐지율이 높다는 것을 의하므로 제안 방식이 더 우수하다고 볼 수 있다. 또한, 비교 방식 2는 제안 방식보다 Precision이 높으나, Recall이 현저히 낮으며 이를 동시에 고려하는 F1 score의 값 또한 훨씬 낮다. 3가지의 비교 방식들과 제안방식을 종합적으로 비교한 결과, 제안 방식이 가장 우수하다고 결론지을 수 있다.

표 5. IoT 네트워크 데이터의 이상탐지 성능 평가 결과
Table 5. The performance of anomaly detection

	Precision	Recall	F1 score	Accuracy	Fall out	Specificity	MCC
Proposed solution (PCA, 1dimension, MMD)	0.982	0.848	0.91	0.86	0.077	0.923	0.638
Comparison 1 (PCA, 1dimension, MD)	0.977	0.650	0.781	0.695	0.077	0.923	0.429
Comparison 2 (NLPCA, 1dimension, MMD)	0.989	0.793	0.880	0820	0.041	0.960	0.594
Comparison 3 (PCA, 12dimension, MMD)	0.954	0.204	0.336	0.328	0.049	0.951	0.150

Ⅴ. 결 론

본 논문은 저사양 IoT 디바이스에도 탑재 가능한 저복잡도 이상탐지 기술을 제안하였다. 사물인터넷의 발전이 급속화 됨에 따라 IoT 디바이스에도 안드로이드와 같은 오픈소스 탑재가 활발히 되고 있다. 그러나 오픈소스의 특성상 네트워크 공격과 같은 외부 공격에 취약하기 때문에, 네트워크 트래픽 데이터를 바탕으로 이상여부를 판단하는 기술 연구가 필수적이다. 제안한 기술은 수집된 정상 데이터에서 PCA를 기반으로 1차원 축소시키는 선형변환벡터를 디자인한 후, IoT 디바이스에 해당 벡터를 탑재하여 실시간으로 들어오는 데이터를 1차원으로 선형변환하여 변형마할라노비스 거리를 기반으로 이상 여부를 탐지하도록 설계되었다. 제안된 기술은 기존의 다른 이상탐지 기술들 대비 낮은 연산량에도 불구하고 우수한 성능을 보임을 안드로이드와 IoT 디바이스에서 수집한 네트워크 데이터를 이용하여 확인하였다. 특히 기존의 마할라노비스 거리를 변형시켜 측정하고자 하는 데이터가 정상 범주로부터 얼마나 떨어져 있는지를 좀 더 정확하게 확인할 수 있게 하여 큰 성능 향상을 이룩하였다. 본 연구 결과를 바탕으로 초저사양 IoT 디바이스에도 저 복잡도 고성능의 이상탐지 시스템을 탑재하여 네트워크 상에서의 안정적인 서비스 제공을 기대한다.

References

[1] K. Lee, B. Kim, and J. Cho, “Design and implementation of security system for providing secure boot and firmware update in low-end IoT device,” *J. KIISE*, vol. 45, no. 4, pp. 321-331, Apr. 2018.

[2] M. Henschke, X. Wei, and X. Zhang, “Data visualization for wireless sensor networks using thingsboard,” *WOCC 2020*, pp. 1-6, Newark, NJ, USA, May 2020.

[3] https://thingsboard.io/

[4] J. Jo, J. Cho, R. Jung, and H. Cha, “IoTivity-Lite: Comprehensive IoT solution in a constrained memory device,” *ICTC 2018*, pp. 1367-1369, Jeju Island, Korea, Oct. 2018.

[5] https://iotivity.org/

[6] S. Behrad, E. Bertin, S. Tuffin, and N. Crespi, “A new scalable authentication and access control mechanism for 5G-based IoT,” *Future Generation Comput. Syst.*, vol. 108, pp. 46-61, Jul. 2020.

[7] https://www.openairinterface.org

[8] C. Seong and K. Rhyu, “Internet of things application service system with open source hardware,” *J. Advanced Marine Eng. and Technol.*, vol. 40, no. 6, pp. 542-547, Jul. 2016.

[9] H. Kye, S. Kim, and M. Kwon, “PCA-based low-complexity anomaly detection for low-end IoT devices,” in *Proc. Symp. KICS*, pp. 402-403, Nov. 2020.

[10] S. Kim, H. Kye, and M. Kwon, “AutoEncoder-based network intrusion detection system,” in *Proc. Symp. KICS*, pp. 404-405, Nov. 2020.

[11] R. U. Islam, M. S. Hossain, and K. Andersson, “A novel anomaly detection algorithm for sensor data under uncertainty,” *Soft Computing*, vol. 22, pp. 1623-1639, Nov. 2018.

[12] M. Hasan, M. M. Islam, M. I. I. Zarif, and M. M. A. Hashem, “Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches,” *Internet of Things*, vol. 7, Sep. 2019.

[13] Y. N. Kunang, S. Nurmaini, D. Stiawan, A. Zarkasi, Firdaus, and Jasmir, “Automatic features extraction using autoencoder in intrusion detection system,” *ICECOS 2018*, pp. 219-224, Pangkal, Indonesia, Oct. 2018.

[14] J. Lee and K. Park, “Network intrusion detection system using feature extraction based on autoencoder in IoT environment,” *KIPS Trans. Softw. and Data Eng.*, vol. 8, no. 12, pp. 483-490, Dec. 2019.

[15] A. Lakhina, M. Crovella, and C. Diot, “Diagnosing network-wide traffic anomalies,” *ACM SIGCOMM 2004: Conf. Comput. Commun.*, vol. 34, no. 4, pp. 219-230, Aug. 2004.

[16] Y. Liu, L. Zhang, and Y. Guan, “Sketch-based streaming PCA algorithm for network-wide traffic anomaly detection,” *ICDCS 2010*, pp. 807-816, Genoa, Italy, Jun. 2010.

[17] J. Camacho, A. Pérez-Villegas, P. García-Teodoro, and G. Maciá-Fernández, “PCA-

based multivariate statistical network monitoring for anomaly detection," *Comput. & Secur.*, vol. 59, pp. 118-137, Jun. 2016.

[18] D. H. Hoang and H. D. Nguyen, "A pca-based method for IoT network traffic anomaly detection," *ICACT 2018*, pp. 381-386, Chuncheon, Korea, Feb. 2018.

[19] A. P. Muniyandi, R. Rajeswari, and R. Rajaram, "Network anomaly detection by cascading k-means clustering and C4.5 decision tree algorithm," in *Proc. Eng.*, vol. 30, pp. 174-182, 2012.

[20] R. Kwitt and U. Hofmann, "Unsupervised anomaly detection in network traffic by means of robust PCA," *ICCGI'07*, Guadeloupe, French Caribbean, Mar. 2007.

[21] Y. Sun, A. Xu, K. Wang, S. Jia, H. Guo, and X. Han, "Anomaly detection of marine diesel engine valve system based on mahalanobis distance," *PHM-Shanghai 2020*, pp. 1-7, Shanghai, China, Oct. 2020.

[22] H. Ya-juan, H. Zhen and S. Guo-fang, "Research for multidimensional systems diagnostic analysis based on improved mahalanobis distance," *2009 IEEE 16th Int. Conf. Ind. Eng. and Eng. Manag.*, pp. 213-217, Beijing, China, Oct. 2009.

[23] Y. Hou, Z. Chen, M. Wu, C. Foo, X. Li, and R. M. Shubair, "Mahalanobis distance based adversarial network for anomaly detection," *ICASSP 2020*, pp. 3192-3196, Barcelona, Spain, May 2020.

[24] H. Om and T. Hazra, "Statistical techniques in anomaly intrusion detection system," *Int. J. Advances in Eng. & Technol.*, vol. 5, no. 1, pp. 387-398, Nov. 2012.

[25] D. Bayarjargal and G. Cho, "Detecting an anomalous traffic attack area based on entropy distribution and mahalanobis distance," *Int. J. Secur. and Its Appl.*, vol. 8, no. 2, pp. 87-94, Mar. 2014.

[26] http://droidcollector.binarytao.com/

[27] D. Cao, S. Wang, Q. Li, Z. Cheny, Q. Yan, L. Peng, and B. Yang, "DroidCollector: A high performance framework for high quality android traffic collection," *2016 IEEE Trustcom/BigDataSE/ISPA*, pp. 1753-1758, Tianjin, China, Aug. 2016.

[28] https://www.kaggle.com/xwolf12/network-traffic-android-malware

[29] L. Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learning Res.*, vol. 9, no. 11, pp. 2579-2605, Nov. 2008.

[30] B. W. Matthews, "Comparison of the predicted and observed secondary structure of T4 phage lysozyme," *Biochimica et Biophysica Acta (BBA) - Protein Structure*, vol. 405, no. 2, pp. 442-451, Oct. 1975.

[31] https://www.kaggle.com/speedwall10/iot-device-network-logs

계 효 선 (Hyoseon Kye)

2018년 3월~현재 : 숭실대학교 전자정보공학부 IT융합전공

2021년 3월~현재 : 숭실대학교 정보통신공학과 학석연계과정

<관심분야> 모바일 네트워크, 이상탐지기술, 인공지능, 연합학습

[ORCID:0000-0001-7808-0387]

권 민 혜 (Minhae Kwon)

2011년 8월 : 이화여자대학교 전자정보통신공학과 학사

2013년 8월 : 이화여자대학교 전자공학과 석사

2017년 8월 : 이화여자대학교 전자전기공학과 박사

2017년 9월~2018년 8월 : 이화여자대학교 전자전기공학과 박사 후 연구원

2018년 9월~2020년 2월 : 미국 Rice University, Electrical and Computer Engineering, Postdoctoral Researcher

2018년 9월~2020년 2월 : 미국 Baylor College of Medicine, Center for Neuroscience and Artificial Intelligence, Postdoctoral Researcher

2020년 3월~현재 : 숭실대학교 전자정보공학부 IT융합전공 조교수

<관심분야> 모바일네트워크, 이상탐지기술, 인공지능, 강화학습, 자율주행

[ORCID:0000-0002-8807-3719]