

전력 시계열 데이터를 이용한 비지도 설비 이상 패턴 탐지 방법

이준희*, 권동우*, 지영민^o

Electric Power Time Series Data-Based Unsupervised Anomaly Detection Method for Equipment

Joonhee Lee*, Dongwoo Kwon*, Youngmin Ji^o

요 약

최근 온실가스, 에너지 규제 강화로 인해 에너지 사용 절감에 대한 필요성이 대두되고 있다. 설비에 이상이 발생하게 되면 불필요한 에너지 소비가 발생하거나, 생산성이 낮아질 수 있다. 따라서 에너지 소비 설비에 대한 이상 탐지의 필요성이 증가하였다. 본 논문은 건물 및 공장 등의 에너지 소비 설비의 실시간 전력 사용량 데이터에 LSTM-오토인코더 기반의 이상 탐지 기법을 적용하여 결과를 분석하고 모니터링한다. 설비에 이상이 발생하게 되면 전력 사용량에 이상 패턴이 나타나게 되고 기존 정상적인 패턴에서 다른 패턴 양상을 보이게 된다. 전력 사용량 데이터에 재구성 기반 이상 탐지 알고리즘을 적용해 실시간으로 에너지 소비 설비의 이상 패턴을 탐지할 수 있음을 확인하였다.

키워드 : EMS, 이상패턴 탐지, 오토인코더, LSTM

Key Words : EMS, Anomaly Detection, AutoEncoder, LSTM

ABSTRACT

Recent greenhouse gas and energy regulations have increased the need to reduce energy use. If an anomaly occurs in the equipment, unnecessary energy consumption may occur or productivity may be lowered. This has increased the need for anomaly detection for energy-consuming equipment. This paper monitor and analyze result applying LSTM-AutoEncoder-based anomaly detection algorithm method for real-time power usage data of energy-consuming equipment in buildings and factories. If an anomaly occurs in the equipment, signs of anomaly appear in the power-usage data, and a pattern different from the normal pattern. By applying the reconstruction-based anomaly detection algorithm to power usage data, we have shown that anomaly patterns in energy-consuming equipment can be detected in real-time.

※ 본 연구는 산업통상자원부(MOTIE)와 한국에너지기술연구원(KETEP)의 지원을 받아 수행한 연구 과제입니다. (No.20212020800120)

◆ First Author : Korea Electronics Technology Institute, joonhee305@keti.re.kr, 정회원

◦ Corresponding Author : Korea Electronics Technology Institute, ym.ji@keti.re.kr, 정회원

* Korea Electronics Technology Institute, dwkwon@keti.re.kr, 정회원

논문번호 : 202403-044-0-SE.R1, Received February 27, 2024; Revised March 25, 2024; Accepted April 1, 2024

I. 서 론

최근 온실가스, 에너지 규제 강화로 인해 FEMS (Factory Energy Management System, 공장 에너지 관리 시스템) 및 BEMS(Building Energy Management System, 건물 에너지 관리 시스템)에 대한 필요성이 대두되고 있다. 건물이나 설비에 문제가 발생하면 그로 인해 불필요한 에너지가 소비되거나 생산성이 저하될 수 있기 때문에 건물 및 설비의 이상 패턴 탐지의 필요성이 높아지고 있다.

이상 패턴 탐지란 설치된 센서들로부터 데이터를 읽어와 정상적인 패턴이 아닌 패턴을 감지하는 것을 말한다. 이상을 사전에 탐지함으로써 건물 및 설비의 고장을 예방하거나 조치할 수 있어 큰 비용 및 자원 손실을 방지할 수 있다. 하지만 모든 설비에 대한 도메인 지식을 알기 어렵고 정상 패턴에 비해 이상 패턴은 거의 관측되지 않거나 없는 데이터 불균형 문제가 있다.

이러한 한계를 극복하기 위해 상관계수, 분포와 같은 통계적 기법, 결정트리, 클러스터링과 같은 머신러닝 기법, 딥러닝 모델 기반의 비지도 학습 등 여러 가지 이상 탐지 방법이 연구되고 있다¹⁻⁴⁾. 특히 센서로부터 수집되는 데이터는 시계열 데이터이기 때문에 RNN (Recurrent Neural Network, 순환 신경망) 계열의 모델을 이용한 여러 방법이 연구되었다.⁵⁻⁷⁾ 그러나 기존 연구에서는 파라미터에 크게 영향을 받거나 데이터의 분포에 따라 결과가 크게 차이가 날 수 있기 때문에 데이터의 형태나 분포에 맞춰 모델을 튜닝해야 하는 한계가 존재한다.

본 논문에서는 LSTM-오토인코더를 이용한 재구성 기반의 이상 탐지 알고리즘⁸⁾을 적용해, 설비에 대한 도메인 지식 없이 에너지 소비 설비의 전력 사용량 데이터를 모니터링하여 이상 패턴을 사전에 탐지할 수 있도록 한 연구 결과를 제시한다.

이후 본 논문의 구성은 다음과 같다. 2장에서는 기존에 진행된 연구의 특징과 한계점, 본 논문에서 진행한 연구의 특징과 장점을 소개한다. 3장과 4장에서는 연구를 진행한 방법과 알고리즘의 소개, 실제 설비에 알고리즘을 적용한 결과를 제시한다. 결론에서는 본 연구의 요약과 개선 사항을 제시한다.

II. 관련 연구

기존 시계열 데이터의 이상 패턴 탐지를 위한 방법으로는 통계적, 머신러닝, 딥러닝 기법 등이 존재한다.¹²⁾ 통계적 기법으로는 데이터의 분포를 분석하고 수집

된 데이터가 분포에서 크게 벗어난 데이터일 때 이상으로 탐지하는 방법이 있다.¹⁰⁾ 단순하고 직관적인 방법이지만 실제 설비에서 가동되는 데이터가 특정한 분포를 따르지 않는다면 적용하기 어렵고, 시간적 특성을 고려하지 못한다는 단점이 존재한다.

회귀 모델을 사용하는 방법으로는 ARIMA(Auto Regressive Integrated Moving Average Model, 자기회귀 누적 이동 평균)라는 회귀 모델을 통해 데이터를 fitting 하고 예측하는 방법이다.¹³⁾ ARIMA는 현재의 값이 과거의 값과 오차에 영향을 받는다는 가정을 바탕으로 시계열 데이터를 모델링 할 수 있다. p 차수를 통해 과거의 어느 시점까지 자기회귀성을 가지는지, q 차수를 통해 오차에 영향을 받는지, d 차수를 통해 정규성을 만들기 위한 차분 횟수를 설정할 수 있다. 위 같은 시계열 모델을 바탕으로 데이터를 모델링 한 뒤 미래값을 예측하여 실제 측정된 값과 비교하고 오차가 클 시 이상으로 판단하는 기법이다. 모델링을 통해 데이터의 시간적 특성을 가지고 값을 예측하는 것이 가능하지만, 차수를 직접 지정해 주어야 하기 때문에 많은 센서가 존재하는 설비에 대해서는 일반화하여 적용하기 어렵다는 단점이 존재한다.

머신러닝 기법으로는 클러스터링 알고리즘이 있다. 클러스터의 개수를 미리 지정하는 k-means(k-평균) 알고리즘과 클러스터의 개수 대신 클러스터의 최소 데이터 개수와 거리를 지정하는 DBSCAN(Density-based spatial clustering of applications with noise, 밀도 기반 클러스터링) 알고리즘¹⁴⁾이 대표적이다. 클러스터링을 통해 데이터 간 밀집도를 파악하고 클러스터와 거리가 먼 데이터를 이상으로 판단하는 방법이다. 특히 DBSCAN은 클러스터의 개수를 지정해 줄 필요가 없어 도메인 지식이 없는 설비에서 유리하다. 하지만 클러스터링 알고리즘 역시 파라미터에 영향을 많이 받기 때문에 센서가 많은 설비에 일반적으로 적용하기 어렵다. 또한 클러스터링 알고리즘은 데이터 간 거리를 계산해야 하므로 시간 복잡도가 크다. 이상치를 탐지하기 위해서는 정상 데이터가 필요한데 데이터가 많아지면 계산 리소스도 함께 증가하기 때문에 실시간 이상 패턴 탐지에는 적합하지 않다.

본 논문에서는 RNN 계열의 LSTM과 오토인코더를 결합한 LSTM-오토인코더를 이용해 이상 패턴 탐지를 진행한다. RNN은 시간의 흐름에 따른 특성을 학습하기 위한 신경망 아키텍처이다. 하지만 시퀀스의 길이가 길어질수록 과거 시점의 정보를 잃어버리는 장기 의존성 문제가 존재한다. 그림 1의 LSTM(Logn-Short Term Memory)은 RNN의 장기 의존성 문제를 해결하기 위한

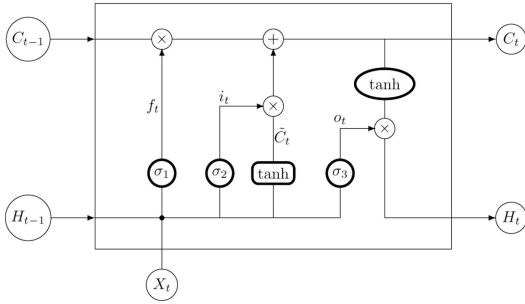


그림 1. LSTM 구조
Fig. 1. LSTM Architecture[5]

아키텍처로 순환 신경망보다 더 오랫동안 과거의 정보를 가지고 학습할 수 있다. 오토인코더는 인코더-디코더 구조로 구성된 딥러닝 모델이며 인코더를 통해 입력된 데이터의 특성들을 압축하고 디코더에서는 압축된 공간벡터를 통해 입력된 데이터를 복원하는 방식인 자기 지도학습(self-supervised) 모델이다. 그림 2의 LSTM-오토인코더는 은닉층의 레이어들을 LSTM으로 구성하여 시간적 특성을 포함하여 학습할 수 있도록 구성된 모델이다.^[5]

재구성 기반 이상 탐지 알고리즘은 그림 3과 같다. 학습을 위한 정상 패턴을 target set이라고 하며 정상 패턴을 입력 데이터로 사용한다. 입력과 출력의 오차를 재구성 오차(Reconstruction Error)라 하며 입력에 대해 재구성 오차가 임계치를 넘지 않는다면 정상, 넘을 경우 이상 패턴으로 판단하는 알고리즘이다.^[7]

LSTM-오토인코더를 이용한 이상 패턴 탐지 알고리즘은 시간적 정보를 포함한 정상 패턴을 학습하는 것이 가능하다. 또한 설계자가 직접 데이터를 분석하고 하이퍼 파라미터를 설정하지 않아도 스스로 학습이 가능하기 때문에 여러 장비들이 동작하는 에너지 소비 설비에서 일반화하여 적용하기 용이하다. 생성한 모델을 바탕으로 빠른 추론이 가능하기 때문에 실시간 이상 탐지에

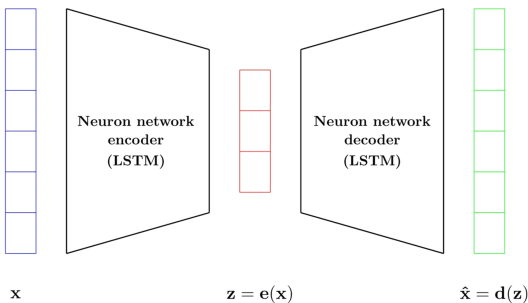


그림 2. LSTM 오토인코더 구조
Fig. 2. LSTM-Autoencoder Architecture[5]

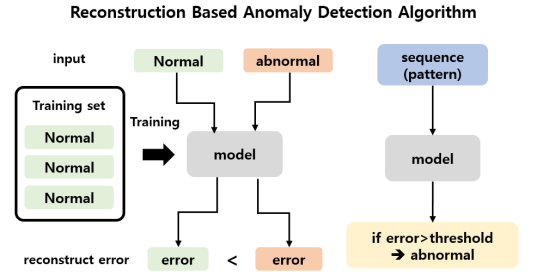


그림 3. 재구성 기반 이상 패턴 탐지 알고리즘
Fig. 3. Reconstruction Based Anomaly Detection Algorithm

유리하다.

본 논문에서는 LSTM 오토인코더를 이용한 재구성 기반의 이상 탐지 알고리즘을 사용하여 실시간 이상 탐지를 적용한다. 도메인 지식이 부족하고, 이상 패턴의 샘플이 부족하거나 없는 상황에서도 정상 패턴에 대해서 스스로 학습하면 재구성 오차를 통해 정상 패턴에 포함되지 않는 이상 패턴을 감지해낼 수 있다. 또한 실시간 이상 탐지에도 용이하기 때문에 해당 알고리즘을 적용해 보았다.

표 1. 이상 탐지를 위한 방법 비교
Table 1. Comparison of methods for anomaly detection

방법	차이점
ARIMA	p, q, d 파라미터에 영향을 많이 받아 여러 장비에 일반화하기 어려움.
DBSCAN Clustering	epsilon, min_samples 파라미터에 영향을 많이 받고, 실시간으로 데이터를 받아왔을 때 전체에 대해 다시 계산을 해야 하기 때문에 실시간 이상 탐지에는 적합하지 않음.
LSTM-AE	하이퍼 파라미터에 대한 영향이 상대적으로 적고, 시계열 데이터의 패턴을 학습할 수 있기 때문에 일반화하여 적용하기 용이함. 다시 학습할 필요 없이 학습된 모델을 가지고 추론을 통해 실시간 이상 탐지가 가능함.

III. LSTM 오토인코더 이상 패턴 탐지

실시간 이상 탐지 시스템은 그림 4와 같다. 설비나 건물에서 발생하는 에너지 데이터를 센서를 통해 수집한다. Ingress-Router와 Query Processor는 각각 데이터의 삽입과 질의를 처리하는 프로세서이며, Ingree-Router를 통해 수집된 데이터는 데이터베이스에 저장된다. Query Processor를 통해 이상 탐지를 위

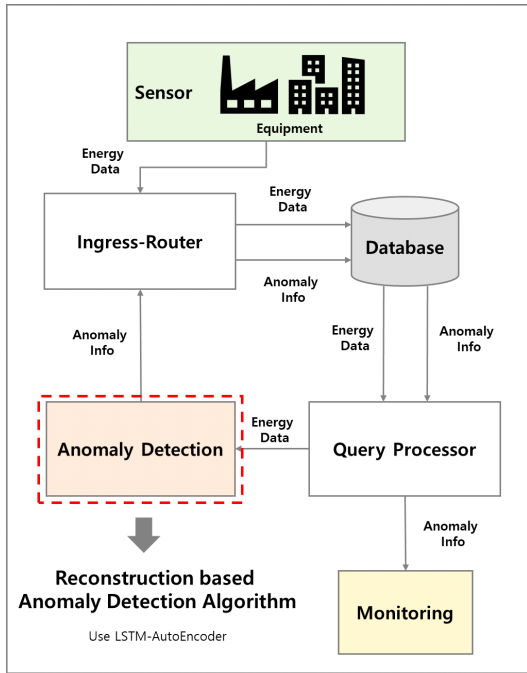


그림 4. 실시간 이상 탐지 시스템 구조
Fig. 4. Real-Time Anomaly Detection System Architecture

한 데이터를 받아와 그림 3의 재구성 기반의 이상 탐지 알고리즘을 적용하고 이상 정보를 계산한다. 이상 정보를 다시 Ingress-Router를 통해 데이터베이스로 저장하고 Query Processor를 통해 이상 정보를 모니터링할 수 있도록 환경을 구축한다.

3.1 학습 데이터 구축

본 논문에서는 실제 운용 중인 공장의 설비 데이터를 이용한다. 학습 데이터는 공정 설비에서 센서로부터 5초에 한 번씩 수집되는 159개 관제점의 전력 사용량 데이터를 사용한다. 관제점은 공정 설비나 장비로부터 데이터를 수집하는 센서를 의미한다.

그림 5는 설비로부터 수집되는 전력 소비 데이터의 일부이다. 관제점 간의 데이터 형태가 달라 관제점마다 특징을 특징하기가 어렵다. 따라서 LSTM을 통해 각 데이터의 특징에 맞게 학습이 가능하도록 한다.

그림 6은 공정 설비의 특정 관제점의 전력 사용량 데이터이다. 관제점 중에서 평소와 다른 패턴을 보이는 구간이 포함되는 관제점을 선정하였다. 3:1 비율로 학습 데이터와 검증 데이터로 분리하였고 테스트 데이터는 이상으로 판단되는 구간을 포함한 2일간의 데이터를 사용하였다.

데이터의 결측치에 대해서는 선형 보간법을 사용하

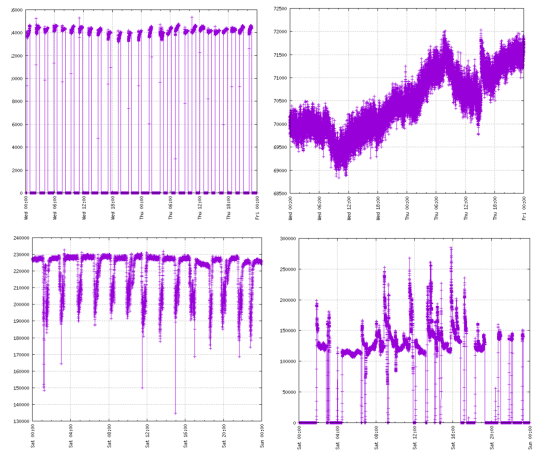


그림 5. 설비의 전력 소비량 데이터
Fig. 5. Power consumption data for equipment

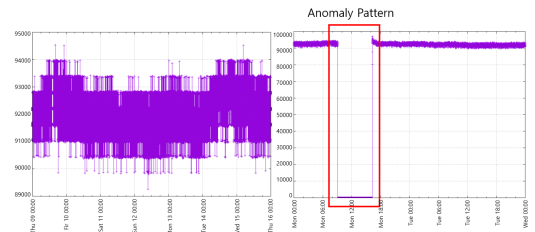


그림 6. (좌) 학습, 검증 데이터, (우) 테스트 데이터
Fig. 6. (Left) train and validation data, (Right) test data

여 처리하였다. 결측치가 없는 데이터에 임의의 5%만큼 결측치를 발생시켜 앞에서 채우기, 뒤에서 채우기, 선형 보간을 진행하고 MAE를 통해 원본 데이터와의

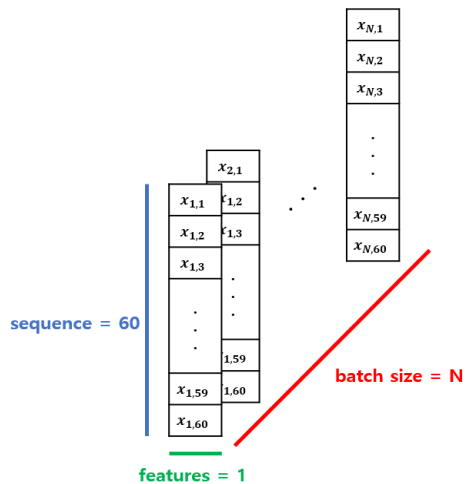


그림 7. 입력데이터 차원
Fig. 7. Dimension of Input data

오차를 비교하였다. 159개 관계점 중 144개의 관계점이 선형 보간의 성능이 높다고 확인되었고 여러 관계점에 보편적으로 사용할 수 있다고 평가하여 결측치 처리는 선형 보간을 사용하였다. 여러 관계점의 데이터 원활한 학습을 위해 z-score 정규화를 이용해 스케일링을 진행하였다. 5분마다 실시간 이상 패턴 탐지를 적용하기 위해 그림 7과 같이 60개 단위로 데이터를 묶어 (batch_size, 60, 1) 차원으로 변환하였다.

3.2 LSTM-오토인코더 모델 구축

학습을 위한 모델의 구조는 인코더와 디코더가 결합된 구조이다. 5분 단위의 패턴을 Input으로 사용하고 Output은 Input을 최대한 유사하게 복원하기 때문에 Input과 Output은 동일하게 (batch_size, 60, 1) 차원으로 구성된다.

3.2.1 모델 아키텍처

그림 8은 LSTM-오토인코더의 레이어 구조이다. 인코더는 전력 사용량 데이터의 시간적 특성을 학습하기 위한 LSTM 레이어 2개와 디코더의 입력 차원을 맞추기 위한 repeat_vector로 이루어져 있으며 디코더는 인코더로부터 생성된 공간벡터를 원본으로 복원하기 위해 인코더와 대칭되는 LSTM 레이어 2개, 입력 데이터의 차원으로 변환하기 위한 time_distributed 레이어로 구성되어 있다.

손실 함수로는 MSE(Mean Square Error, 평균 제곱 오차)를 사용하고, 하나의 패턴에 대한 재구성 오차는 다음 수식 1과 같이 계산한다. 재구성 된 값과 실제 값의 차이를 제공한 뒤 N으로 나누어 평균을 구하여 재구

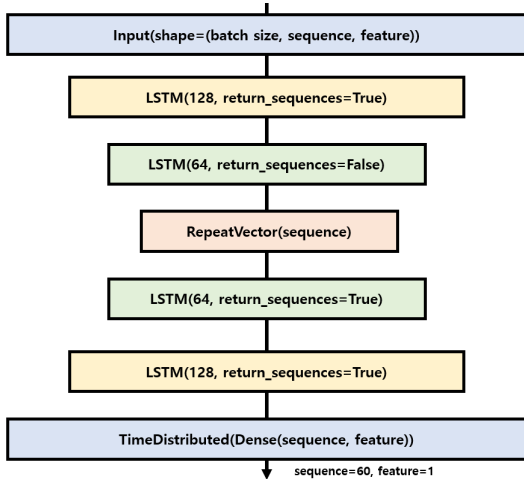


그림 8. LSTM-오토인코더 레이어 구조
Fig. 8. LSTM-Autoencoder Layer Architecture

성 오차로 사용한다. 실시간 이상 패턴 탐지에서도 수식 1을 통해 재구성 오차를 계산하고 임계치와 비교하여 이상 여부를 식별한다.

$$MSE_{recon} = \frac{1}{N} \sum_{i=1}^N (\bar{x}_i - x_i)^2 \quad (1)$$

N = 패턴(시퀀스) 길이 \bar{x} = 재구성 값 x = 원본 값

3.2.2 모델 학습 및 임계치 선정

학습 데이터와 LSTM-오토인코더를 이용해 그림 9 처럼 재구성 오차가 작아지는 방향으로 학습하게 된다. 또한 재구성 오차를 기반으로 임계치를 결정한다. epoch는 100으로 설정하고 early stop을 기법을 적용해 과적합을 방지한다. val_loss가 가장 낮은 모델을 선정해 실시간 이상 탐지에 적용한다.

임계치는 이상 패턴이 정상 패턴에 비해 오차가 크다는 점을 이용하여 결정한다. 임계치는 학습하는 과정에

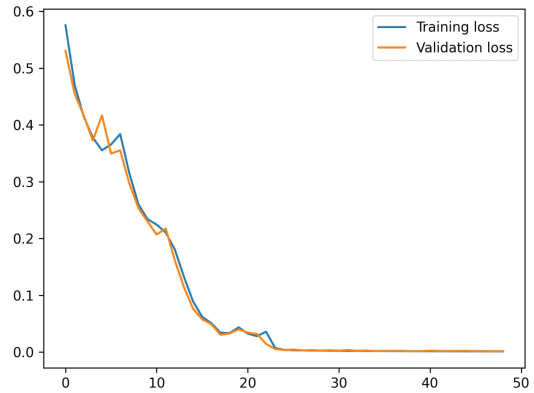


그림 9. 학습 데이터 및 검증 데이터의 오차
Fig. 9. Loss of training and validation data

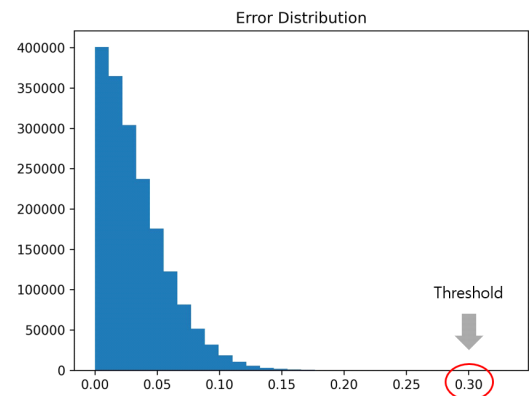


그림 10. 학습 데이터의 오차 분포
Fig. 10. Error distribution of training data

서 선정된 모델에 대해 학습 데이터에 대한 재구성 오차를 구한 뒤, 그림 10처럼 재구성 오차 중 최댓값을 임계치로 사용한다. 정상 패턴의 재구성 오차 범위를 넘어갈 경우 이상으로 판단할 수 있다.

3.3 알고리즘

알고리즘 1은 5분마다 실시간으로 이상 패턴을 탐지하는 알고리즘이다. metric은 특정 설비를 뜻하며, 설비마다 알고리즘을 한 번씩 적용한다. 현재 시간부터 5분 이전의 구간 사이의 데이터를 질의한다. standard scaler를 통해 z-score 정규화를 진행한다. 정규화된 데이터를 모델에 추론시키고 재구성 오차를 구한 뒤 임계치와 비교하여 정상인지 이상인지를 판단한다.

```

Algorithm - Real-Time Anomaly Detection
while(True):
    for metric in metrics:
        1. Input:  $x_m = \text{query}(t, t-5m)$             $x_m = Dm(60 \times 1)$ 
        2. Scaled:
            $x'_m = \frac{x_m - \mu}{\sigma}$ ,            $x'_m = Dm(1 \times 60 \times 1)$ 
            $x'_m = x'_m.reshape(1 \times 60 \times 1)$ 
        3. output:
            $y_m = \text{MODEL}(x'_m)$             $y_m = Dm(1 \times 60 \times 1)$ 
        4. loss:  $l_m = \text{MSE}(y_m)$             $l_m = float$ 
        5. if loss > threshold then
        6.     status = 1           1(Abnormal)
        7. else:
        8.     status = 0           0(Normal)
        9.     t += 5m
    
```

알고리즘 1. 실시간 이상 탐지 알고리즘
Algorithm. 1. Real-Time Anomaly Detection Algorithm

IV. 실험 결과

테스트 데이터 및 실제 구동 중인 설비에 알고리즘을 적용한 결과를 분석한다. 4.1절에서는 정상 패턴을 학습한 후 학습된 패턴이 아닌 패턴에 대해서 이상으로 판단하는지에 대한 여부, 학습된 패턴에 대해서 정상으로 판단하는지를 실험한다. 4.2절에서는 선정한 관제점 이외의 다른 관제점들에 대해서 일괄 적용하여 실시간 이상 탐지가 가능한지에 대한 여부와 실증 서버에서 실시간 이상 탐지의 성능을 실험한다.

4.1 테스트 데이터 적용 결과

테스트 데이터에 대한 적용 결과는 그림 11와 같다. Original Value은 실제 센서로부터 수집된 데이터이고 Reconstruction Value는 LSTM-AE를 통해 재구성된 값이다. Reconstruction Error은 재구성 값과 원본 값 간의 MSE이며 Threshold(임계치)와 비교하여 임계치보다 클 경우 Anomaly가 1, 낮다면 0이 된다. 그림 6의 학습 데이터의 패턴을 적절히 학습하여 정상 패턴에 대해서는 제대로 복원을 하고 표시된 이상 구간에 대해서는 복원을 하지 못한 것을 확인할 수 있다. 복원을 하지 못한 구간에 대해서는 재구성 오차가 임계치를 넘어 이상으로 판단된 것을 확인할 수 있다.

또한 서로 상관관계가 높은 설비의 경우 특정 관제점에서 이상 패턴이 나타날 때 다른 관제점에도 이상 패턴이 나타날 가능성이 존재한다. 그림 12는 상관관계가 높은 장비들의 개별 데이터이고, 그림 13은 해당 장비들에서 이상으로 판단된 Anomaly 값의 합계이다. 특정

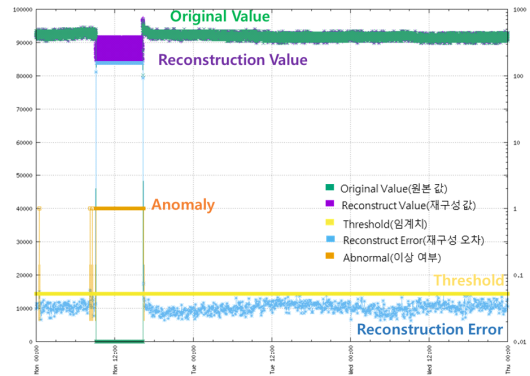


그림 11. 테스트 데이터의 이상 탐지 알고리즘 적용 결과
Fig. 11. Result of applying the anomaly detection algorithm on test data

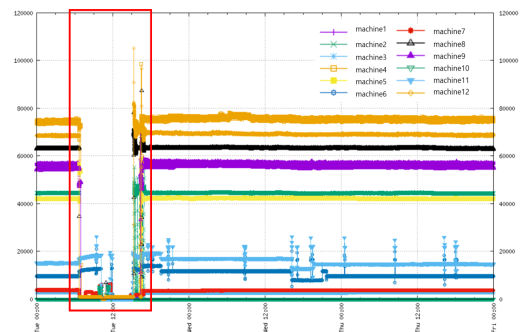


그림 12. 설비 내 장비 간 상관관계가 있는 케이스
Fig. 12. Cases where Machines is correlated in Equipment

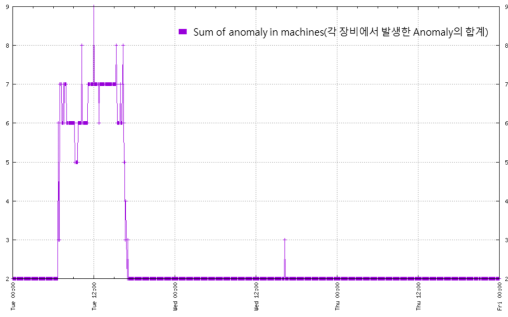


그림 13. 관계성이 높은 장비들에서 발생한 이상 분포
Fig. 13. Anomaly distribution of Cases where Machines is correlated in Equipment

구간에서 이상으로 판단된 빈도가 높아짐을 확인할 수 있다.

4.2 설비 실시간 이상탐지 알고리즘 적용 결과

선정한 관제점에 대해 실시간 이상 탐지를 적용한 결과는 그림 14와 같다. 2023년 10월 16일 하루의 모니터링 결과이다. 일부 구간을 제외한 대부분의 구간에선 정상으로 판단된다.

그림 14에서 이상으로 판단된 구간을 확대하면 그림 15와 같다. 해당 구간에 대해 분석하면 임계치와 재구성 오차의 근소한 차이로 이상으로 판단됨을 확인할 수 있다. 학습 데이터로부터 임계치를 결정하는데 모델이 학습 데이터에 적합되어 있기 때문에 학습 데이터 이외에 다른 데이터에선 상대적으로 오차가 증가할 가능성이 존재한다.

그림 16은 테스트 한 관제점 이외에 다른 관제점들에 실시간 이상 탐지 알고리즘을 적용한 결과이다. 서로 다른 형태의 데이터 패턴에 대해서도 잘 학습하여 정상으로 판단되는 것을 확인할 수 있다.

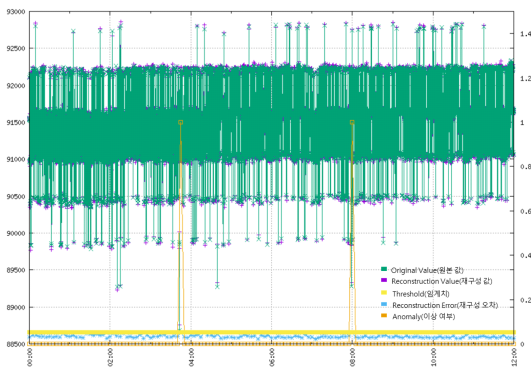


그림 14. 설비의 실시간 이상 탐지 알고리즘 적용 결과
Fig. 14. Results of applying real-time anomaly detection algorithms to sensor data

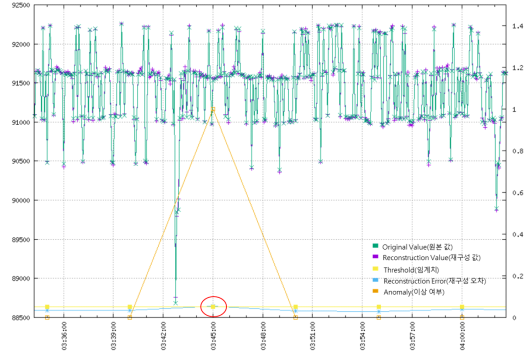


그림 15. 재구성 오차와 임계치의 차이가 작은 케이스
Fig. 15. Case the difference is small where between reconstruction error and threshold

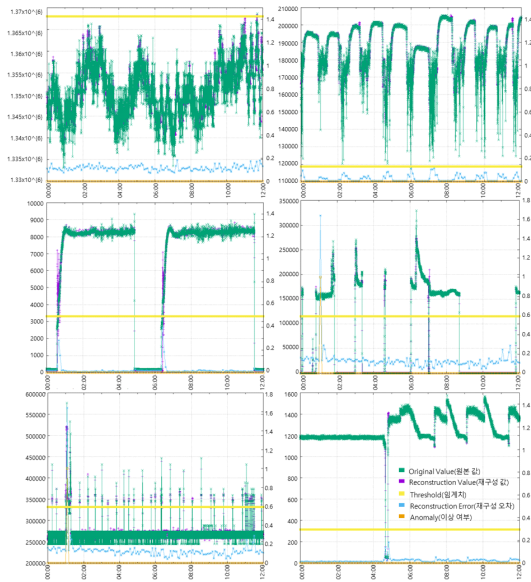


그림 16. 다른 관제점에 실시간 이상 탐지 알고리즘 적용 결과
Fig. 16. Results of applying real-time anomaly detection algorithms to another sensor data

그림 17은 실증 서버에서 이상 패턴 탐지를 적용한 로그 화면이다. 하나의 관제점에 대해 모델 추론에 소요되는 시간은 약 0.4초 내외로 측정되었고 160개의 관제점에 대해 데이터 질의시간을 포함하여 약 3분 내외의 실행 간격으로 이상 탐지를 적용할 수 있음을 확인하였다.

```

send : reconstruction data ---> 0.377563sec
=> Nan 개수 : 0
Send Success : 200
send : reconstruction data ---> 0.383814sec
=> Nan 개수 : 0
Send Success : 200
send : reconstruction data ---> 0.390268sec
=> Nan 개수 : 0
Send Success : 200
send : reconstruction data ---> 0.379604sec
=> Nan 개수 : 0
Send Success : 200
send : reconstruction data ---> 0.381967sec
=> Nan 개수 : 0
Send Success : 200
send : reconstruction data ---> 0.381599sec
=> Nan 개수 : 0
Send Success : 200
send : reconstruction data ---> 0.381183sec
=> Nan 개수 : 0
Send Success : 200
send : reconstruction data ---> 0.374347sec
=> Nan 개수 : 0
Send Success : 200
    
```

그림 17. 서버에서 가동중인 실시간 이상 탐지
 Fig. 17. Real-time Anomaly detection run-time log on the server

V. 결론

최근 에너지, 온실가스 규제로 인해 에너지 사용 절감에 대한 필요성이 대두되고 있다. 건물 및 공장 설비에서 이상이 발생할 경우 불필요한 에너지 소모가 발생하거나 생산성이 저하될 수 있기 때문에 이상 패턴을 감지하는 것이 중요하다. 하지만 모든 설비에 대해 정보를 얻기 어렵고 정상 패턴에 비해 이상 패턴의 비중이 적은 데이터 불균형 문제가 있다. 따라서 본 논문에서는 에너지 소비 설비에 대한 설비 정보와 도메인 지식 없이 Data-Driven으로 이상 패턴을 분석하는 방법을 적용해 보았다. LSTM-오토인코더를 사용한 재구성 기반의 이상 탐지 알고리즘을 통해 정상 패턴들을 학습하고 재구성 오차를 통해 정상 패턴과 다른 패턴을 식별하는 것이 가능하다고 판단 되었다. 특정한 형태의 데이터뿐만 아니라 서로 다른 형태의 데이터에도 적용이 가능하다. 도메인 지식이 부족한 상태에서 알고리즘을 적용하기 때문에 정확한 성능 평가 지표를 선정하기 어렵다. 하지만 알고리즘을 적용하여 얻을 수 있는 이상 패턴의 정보와 각 설비 환경에 맞는 데이터 분석을 통해 구체적인 성능 평가 지표를 선정하고 정확한 이상 패턴 탐지를 적용할 수 있을 것으로 기대된다. 추후 Anomaly Transformer^[8] 이나 CNN^[10]을 추가로 결합한 재구성 기반 모델을 통해 성능 비교 및 향상을 위한 연구를 진행할 계획이다.

References

[1] Y. Ji and D. Kwon, "A study on the method of diagnosing abnormalities in energy-using

facilities through correlation analysis," in *Proc. Symp. KICS*, vol. 2022 no. 11, pp. 82-83, Nov. 2022.

[2] A. A. Cook, G. Mısırlı, and Z. Fan, "Anomaly detection for IoT time-series data: A survey," in *IEEE Internet of Things J.*, vol. 7, no. 7, pp. 6481-6494, Jul. 2020. (<https://doi.org/10.1109/JIOT.2019.2958185>)

[3] J.-S. Chou and A. S. Telaga, "Real-time detection of anomalous power consumption," *Renewable and Sustainable Energy Rev.*, ISSN 1364-0321, vol. 33, pp. 400-411, May 2014. (<https://doi.org/10.1016/j.rser.2014.01.088>)

[4] D. Deng, "Research on anomaly detection method based on DBSCAN clustering algorithm," *2020 5th Int. Conf. ISCTT*, pp. 439-442, Shenyang, China, Nov. 2020. (<https://doi.org/10.1109/ISCTT51595.2020.00083>)

[5] H. D. Nguyen, K. P. Tran, S. Thomassey, and M. Hamad, "Forecasting and anomaly detection approaches using LSTM and LSTM Autoencoder techniques with the applications in supply chain management," *Int. J. Inf. Manag.*, ISSN 0268-4012, vol. 57, pp. 102282, Apr. 2021. (<https://doi.org/10.1016/j.ijinfomgt.2020.102282>)

[6] O. I. Provotar, Y. M. Linder, and M. M. Veres, "Unsupervised anomaly detection in time series using LSTM-based autoencoders," *2019 IEEE Int. Conf. ATIT*, pp. 513-517, Kyiv, Ukraine, Dec. 2019. (<https://doi.org/10.1109/ATIT49449.2019.9030505>)

[7] H.-S. Nam, Y.-K. Jeong, and J. W. Park, "An anomaly detection scheme based on LSTM autoencoder for energy management," *2020 Int. Conf. ICTC*, pp. 1445-1447, Jeju Korea, Dec. 2020. (<https://doi.org/10.1109/ICTC49870.2020.9289226>)

[8] J. Xu, et al., "Anomaly transformer: Time series anomaly detection with association discrepancy," *arXiv preprint arXiv:2110.02642*,

2022.

(<https://doi.org/10.48550/arXiv.2110.02642>)

- [9] M. Markou and S. Singh, "Novelty detection: A review –part 1: Statistical approaches," *Signal Processing*, ISSN 0165-1684, vol. 83, no. 12, pp. 2481-2497, Dec. 2003.

(<https://doi.org/10.1016/j.sigpro.2003.07.018>)

- [10] S. Hong and H. Kim, "Integrated CNN for specified object filtering and abnormal motion detection for smart factory," *J. KICS*, vol. 48, no. 8, pp. 1001-1011, Aug. 2023.

(<https://doi.org/10.7840/kics.2023.48.8.1001>)

이 준 희 (Joonhee Lee)



2024년 : 성결대학교 공학사
2023년~현재 : 한국전자기술연
구원 에너지IT융합연구센터
연구원

<관심분야> 인공지능, 빅데이
터, 데이터 플랫폼

권 동 우 (Dongwoo Kwon)



2010년 : 계명대학교 공학사
2012년 : 계명대학교 공학석사
2017년 : 계명대학교 공학박사
2018년 : 계명대학교 산업기술연
구소 연구원
2018년~현재 : 한국전자기술연
구원 에너지IT융합연구센터
책임연구원

<관심분야> IoT, 데이터 플랫폼, 에너지데이터 분석

지 영 민 (Youngmin Ji)



2005년 : 경희대학교 공학사
2007년 : 고려대학교 공학석사
2007년~2009년 : 모토로라 Things
to Things Research Center
연구원
2010년~2012년 : 삼성SDS 정보
통신기술연구소 선임 연구원

2012년~현재 : 한국전자기술연구원 에너지IT융합연구
센터 책임연구원

<관심분야> IoT, 빅데이터, 인공지능 기반 상황 인지
분석