

<http://dx.doi.org/10.17703/JCCT.2024.10.1.453>

JCCT 2024-1-53

## 수집 데이터 기반 경량 이상 데이터 감지 알람 시스템 개발

### Evaluation of Edge-Based Data Collection System through Time Series Data Optimization Techniques and Universal Benchmark Development

조우진\*, 구재회\*\*

Woojin Cho\*, Jae-hoi Gu\*\*

**요약** 전 세계적으로 기후 위기와 에너지 비용 증가 등의 문제로 인해 에너지 절감과 관리에 대한 관심이 증대되고 있다. 대한민국의 경우 전체 에너지 사용량의 약 53.5%가 산업 단지에서 사용하여 이와 관련한 에너지 절감 포인트를 찾고자 유사한 에너지 유틸리티를 사용하는 기업 간의 "공유 네트워크 유틸리티 플랜트"를 통해 문제점을 개선하고자 하였다. 이러한 에너지 절감을 위해서 활용하는 다양한 기법들과 공장의 안정적인 운영을 위해서는 데이터의 안정적 공급이 그 무엇보다 중요하다. 그를 위해 데이터를 안정적으로 공급할 수 있는지에 대해 알아볼 수 있는 이상 데이터 감지 시스템과 알람 시스템의 대다수는 에너지 관리 시스템에 종속되어 한계가 있었고 에너지 관리 시스템의 구축은 대단위 시스템의 구축으로 공간, 에너지적 한계가 있는 소형 공장에서 구축이 어려웠다. 본 논문에서는 문제점들을 극복하고자 적은 공간과 전력을 소비하는 임베디드 디바이스에 데이터 수집 시스템과 이상 데이터 감지 알람 시스템을 구축하고, 데이터 수집을 하는 보편적인 기관에서 이상 데이터 감지 알람 시스템의 활용 가능성과 구축 과정에 대한 연구를 수행하였다.

**주요어** : 알람 시스템, 데이터 수집 시스템, 데이터베이스, 시계열 데이터, 이상 데이터 감지, 구축

**Abstract** Due to global issues such as climate crisis and rising energy costs, there is an increasing focus on energy conservation and management. In the case of South Korea, approximately 53.5% of the total energy consumption comes from industrial complexes. In order to address this, we aimed to improve issues through the 'Shared Network Utility Plant' among companies using similar energy utilities to find energy-saving points. For effective energy conservation, various techniques are utilized, and stable data supply is crucial for the reliable operation of factories. Many anomaly detection and alert systems for checking the stability of data supply were dependent on Energy Management Systems (EMS), which had limitations. The construction of an EMS involves large-scale systems, making it difficult to implement in small factories with spatial and energy constraints. In this paper, we aim to overcome these challenges by constructing a data collection system and anomaly detection alert system on embedded devices that consume minimal space and power. We explore the possibilities of utilizing anomaly detection alert systems in typical institutions for data collection and study the construction process.

**Key words** : Notification System, Data Collection System, Database, Time series data, Anomaly data detection, Established

\*정회원, 고등기술연구원 연구원 (제1저자)  
\*\*정회원, 고등기술연구원 연구위원 (교신저자)  
접수일: 2023년 10월 16일, 수정완료일: 2023년 11월 9일  
게재확정일: 2023년 11월 10일

Received: October 16, 2023 / Revised: November 9, 2023  
Accepted: November 10, 2023  
\*\*Corresponding Author: jaehoi@iae.re.kr  
Institute for Advanced Engineering

## I. 서론

전 세계적으로 에너지의 비용 증가와 여러 위기 속 에너지 절감과 에너지 관리에 대한 관심이 증대되고 있다. 대한민국의 에너지 현황은 산업단지가 매우 높은 비중을 차지하고 있는데, 산업단지는 2018년 기준으로 국가 전체 에너지 사용량의 약 53.5%의 에너지를 연간 사용한다[1]. 이러한 산업단지의 경우 산업단지 내에 유사 업종을 하는 공장이 모여 있는 특성을 지니고 있어 공장마다 사용하는 유틸리티가 유사한 경우가 많다. 하지만 대다수의 공장에서는 유사한 유틸리티를 사용하는 이점을 살리지 못하고 필요로 하는 에너지를 개별 계약을 맺거나, 자체 생산을 하여 사용하는 경우가 대다수이다. 이러한 에너지 유틸리티 사용법은 자체 생산 시 과잉 생산분 혹은 에너지가 부족할 시 발생하는 문제점을 대비할 수 없으며, 높은 단가의 계약을 통해서만 거래가 가능하다는 단점 등을 지니고 있다.

이러한 유사한 에너지 유틸리티를 사용하는 이점을 살리고 문제점을 극복하고자 제안된 개념이 “산업단지의 에너지 공동 네트워크 플랜트” 개념이다. 이와 같은 개념은 공용 설비와 같은 안정적 공급망을 통해 공통 에너지 유틸리티 활용에 안정성을 갖추고, 에너지 생산 기업 수요 기업 간의 거래를 수요, 공급 예측 및 라우팅 알고리즘을 통한 가이드선 제시로 에너지 절감과 에너지의 낭비를 최소화하는 개념이다[2][3]. 이러한 에너지 유틸리티를 통해 산업단지 전체의 에너지 사용량 절감 및 불필요한 낭비를 막고자 하는 연구를 진행하고 있다[4].

공장 및 산업단지에서는 수많은 센서가 존재하는데 그중 본 연구와 밀접하게 연관되어있는 에너지 사용과 같은 측정은 실제 돈과 직결되며 센서 데이터의 이상은 매우 중대한 문제를 일으킬 가능성이 높다. 어떠한 센서나 측정 장비도 완벽할 수는 없으며 설비 또한 제대로 동작하지 않을 수 있다. 단순 센서의 문제가 아니라 실제 장비 혹은 설비 등의 이상으로 인해 이상치의 값이 나와 문제가 발생할 시 재물적 피해뿐만이 아닌 인명피해까지 발생할 수 있다. 따라서 이상 데이터를 탐지하는 것은 공장과 같은 시설을 운영 함에 있어 매우 중요한 역할을 한다. 데이터의 이상에 대해 인지하지 못한다면, 공장의 운전, 경제적 문제 등 수많은 문제를 유발할 수 있다.

따라서 본 논문에서는 데이터 수집 시스템에서 수집되는 데이터의 이상이 발견될 시 특정 매체를 통하여 알림을 전송하는 시스템을 구축한다. 이상 데이터 감지 알림 시스템은 이상 데이터 감지 시스템과 알림 시스템 두 가지의 시스템을 지니고 있다. 또한 데이터의 이상을 감지할 수 있도록 실제 구동되는 공장에서 수집되는 데이터를 감시하여 통계적 기법을 활용하여 구분하며, 센서의 이상 유무, 데이터 수집 불가 등 특수 상황에 따라 알림을 보내 상태를 확인한다. 또한 에너지 절감을 위해 구축하는 시스템을 기존의 몇천 와트(Watt)가 넘는 전력을 사용하는 서버 시스템에 구축하는 것은 에너지 절감과는 상충되며, 산업단지 내 공장의 환경 특성상 기존의 서버실을 관리 및 구성하는 것은 쉽지 않다. 따라서 본 논문에서는 이상 감지 알림 시스템을 기존의 서버 시스템에서의 구축으로 진행하는 것이 아닌 저사양의 저전력 임베디드 디바이스를 통해 데이터 수집 시스템과 이상 데이터 감지 시스템을 함께 구축하여 적은 공간 및 적은 전력으로도 구축 가능한 시스템을 확인한다[5].

본 논문은 저전력 저사양 디바이스에서의 이상 데이터 감지 및 알림 시스템을 개발하고 구축하며 최적화를 진행해 가상유틸리티 플랜트 네트워크 시스템의 안정화에 기여할 수 있다.

## II. 배경 및 관련 연구

### 1. 이상 감지

이상 감지 혹은 이상 징후 감지라고 하는 기법은 정상적 데이터 패턴에서 벗어난 사례를 식별하는 데 주로 사용되는 데이터 분석적 기법 중 하나이다. 이상 감지 기법은 사용자 행동 감지, 거래 모니터링, 센서 데이터 분석, 서버 로그 분석 등 사회의 여러 분야에서 사용하고 있다. 하지만 이러한 기법은 공장이라는 특성에 의해 보다 중요한 요소가 된다[6]. 기존의 연구들은 대부분 이상을 감지하는 인공지능적 방법론의 제시만을 주로 진행하였다. 하지만 저사양 임베디드 디바이스에서 인공지능 활용 이상 데이터 감지 기법은 높은 부하에 의해 어려운 것이 실정이며, 이상 데이터를 감지 시 실제로 알림을 주는 시스템에 대한 연구는 부족한 실정이다.

## 2. 데이터 수집 시스템

데이터 수집 시스템이란 데이터를 수집하는 센서에서부터 수집된 데이터를 저장되는 전 과정과 데이터베이스에 적재되는 과정 및 전처리 과정 모두를 통칭하는 것을 데이터 수집 시스템이라고 한다[7]. 기존의 데이터 수집 시스템은 DAQ(Data Acquisition)라고 하는 센서에서 전기적, 화학적 신호등을 측정하여 데이터로써 저장하는 시스템만을 의미하였으나, 빅데이터와 인공지능 등의 광범위한 데이터를 다루는 개념의 등장으로 인해 데이터 수집의 전 과정을 아우르는 개념이 되었다.

## 3. 알람 시스템

알람 시스템은 다양한 분야에서 특정 패턴을 벗어나는 문제 등의 문제가 발생하였을 때, 이상 패턴이 발생했다는 알람을 주는 시스템을 통칭한다.

기존의 알람 시스템에 대한 연구는 대체로 노인, 어린이에 관련된 연구나 스마트 모빌리티 관련 연구가 주로 이루어졌다. 노인과 어린이의 특수 패턴을 분석하여 특정 패턴이 발견될 시 알람이 가게 하거나, 스마트 모빌리티의 상태 추적에 관한 연구들이 주로 이루어졌다[8][9].

에너지 관련 분야에서는 관련 부분이 에너지 관리 시스템 내의 알람 시스템에 국한되어있는 경우가 대다수였다. 에너지 관리 시스템까지는 구축하지 않은 소형 기업에서 사용하기란 매우 어려운 부분이 있다[10]. 따라서 본 연구에서는 단순히 공장뿐 아니라 센서로 데이터를 수집하는 대다수의 환경에서 사용할 수 있는 보편적인 이상 데이터 감지 알람 시스템 제안하는 연구를 진행했다.

## III. 데이터 수집 시스템

이상 데이터 감지 알람 시스템은 수집되는 데이터를 지속해서 감시하여 이상 데이터가 있는지 탐지한 후 이상이 있다면 알람을 주는 역할을 한다. 따라서 이상 데이터 감지 알람 시스템에 앞서 데이터를 수집하는 데이터 수집 시스템이 필수적으로 선행되어야 한다. 또한 기존 공장의 열악한 상황에 한가지 임베디드 디바이스에서 데이터 수집 시스템 및 알람 시스템을 모두 구축

하기 위해서는 센서 데이터 수집에 최적화가 되도록 데이터 수집 시스템을 최적화하여 적은 자원을 효율적으로 활용할 수 있도록 구축하여야 한다.

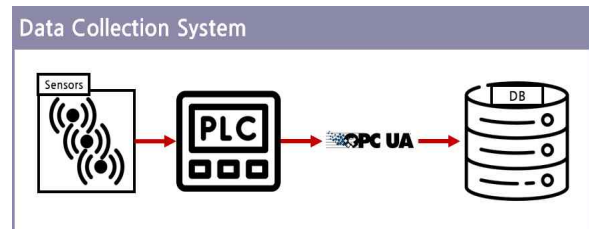


그림 1. 데이터 수집 시스템 구조 개요  
Figure 1. Data Collection System Architecture Overview

데이터 수집 시스템은 그림 1과 같이 구성이 되어있다. 실시간으로 수집되는 센서의 데이터는 PLC와 OPC-UA Server를 거쳐 데이터베이스에 실시간으로 적재 된다[11]. 데이터베이스는 데이터 감시를 위해 읽기에 유리한 관계형 데이터베이스의 형식을 띠는 것이 실험 결과 가장 좋은 처리량을 보여주었기 때문에 MySQL을 사용하는 것으로 진행하였다. 하지만 기본 MySQL은 InnoDB라는 B-tree 스토리지 엔진을 사용하는 문제에 의해 높은 쓰기 부하가 발생하여 데이터 수집 자체에 문제가 발생할 수 있다[12]. 따라서 높은 쓰기 부하를 발생시키는 시계열 데이터 쓰기에 최적화할 수 있도록 Key-value store인 RocksDB를 Backend 스토리지 엔진으로 활용하는 MyRocks를 데이터 수집 시스템으로 선택하여 수집을 진행한다[13].

실증 중인 공장에서 단순히 수집만을 진행하는 데이터 수집 시스템을 Raspberry PI 4 Model B 4GB 모델에 적용하였을 때 100개가량의 센서 데이터가 1초 단위로 삽입이 되어도 CPU는 평균 30%대, RAM은 평균 2GB 이하를 사용하여 자원은 매우 적게 사용하였으며, 이를 실증 공장에서 약 100일 동안 실 운전하였을 때도 데이터 수집과 관련하여 문제가 발생하지 않는 것을 확인하였다[14].

데이터 수집 시스템에 사용한 MyRocks 버전은 MySQL 5.6.35기반의 RocksDB 8.2.0을 지원하는 MyRocks를 사용하였다[15].

#### IV. 이상 감지 알림 시스템 구축

##### 1. 이상 감지 알림 시스템

본 연구에서 구축한 이상 감지 알림 시스템의 경우 데이터 수집 시스템에서 실시간으로 수집되는 센서 데이터가 데이터베이스에 적재 되면, 알림 시스템은 지속적으로 데이터를 감시하다 이상 데이터 징후를 발견하면 사용자에게 알림을 전달하는 시스템이다.

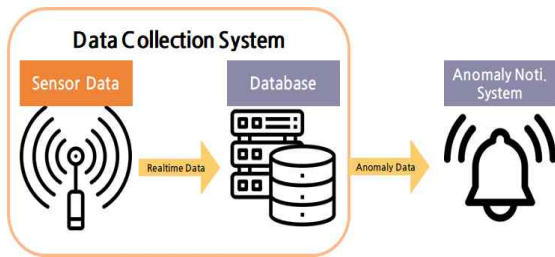


그림 2. 이상 감지 알림 시스템 구조 개요  
Figure 2. Anomaly Detection Notification System Structure Overview

이상 감지 알림 시스템 구조의 개략적인 구성은 그림 2와 같다. 데이터 수집 시스템을 통해 센서에서 실시간으로 데이터를 수집한다. 데이터 수집 시스템에서는 MySQL을 사용했으나 Backend 스토리지 엔진을 Key-value store를 사용하여 높은 쓰기 부하를 처리할 수 있는 데이터 수집 시스템에서 데이터 수집에 문제가 발생하는 특정 상황에 대한 전처리를 진행한 후 이상 데이터 감지 알림 시스템에 전달한다.

데이터베이스에서 데이터 판별을 진행하기 때문에 쓰기가 완료되기 전 읽을 수 없어 그를 방지하고자 특정 단위 시간을 기간으로 두고 데이터를 읽는 것으로 문제를 방지하였다.

해당 이상 데이터 감지 알림 시스템은 크게 두 부분으로 분류된다. 실시간으로 데이터베이스 적재되는 실시간 센서 데이터의 데이터 스트림을 감시하는 이상 데이터 감지 시스템 부분이 한 부분이다. 이상 데이터 감지 시스템이 이상 데이터로 판별하였을 시 이상 데이터 감지 시스템에서 얻은 결과에 따라 사용자에게 알림을 제시하는 알림 시스템 두 가지 시스템이 이상 감지 알림 시스템을 구성한다. 임베디드 디바이스에 데이터 수집 시스템과 이상 데이터 감지 알림 시스템을 함께 구축하여 활용하는 과정에 대해 알아본다.

##### 1) 이상 데이터 감지 시스템

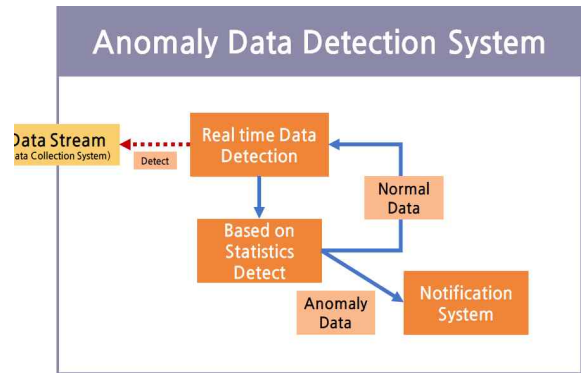


그림 3. 이상 감지 알림 시스템 구조 개요  
Figure 3. Anomaly Detection Notification System Structure Overview

이상 데이터 감지 시스템은 그림 3과 같이 구성된다. 데이터 수집 시스템에서 수집한 데이터가 데이터베이스에 적재 된 후 단위 시간 전의 데이터 스트림을 실시간 데이터 감지 기법을 통해 모든 데이터를 지속적으로 확인한다. 이때 실시간 데이터 스트림의 데이터를 감지할 수 있으나, 데이터베이스에서 데이터를 감지하는 이유는 데이터베이스에 데이터를 적재하기 전의 전처리 기법 중 일부를 해당 데이터 감지 시스템에서 별도로 활용할 시 낮은 성능의 임베디드 디바이스에 의해 한계가 있을 수 있어 중복 작업 제거하기 위함과 데이터베이스에 쓰이기 전에 읽어 발생하는 문제를 방지하기 위해 데이터베이스에서 데이터를 감지하는 기법을 활용하였다.

데이터를 탐지한 후 데이터의 이상을 감지하기 위해서 현재 실증을 진행하는 공장에서 기존에 수집되었던 데이터를 기준으로 다양한 오류나 이상치를 판별할 수 있는 통계 기반 감지 기법을 활용하여 데이터 이상을 판별한다.

데이터베이스에 적재 되는 데이터는 1초 단위로 수집되는 데이터이기 때문에 초 단위로 수집되는 모든 데이터를 확인할 수 있도록 데이터 감지 및 판별 부하를 최소화할 수 있는 기법에 대한 연구를 추가로 진행하였다. 그 결과 데이터베이스 적재 및 조회 오버헤드가 가장 큰 것을 확인하여 데이터베이스 최적화를 진행하였다.

데이터를 매초 확인한 후 정상 데이터로 판별되면 기존 데이터 감지 시스템으로 돌아가 지속적인 데이터 탐지를 진행한다. 이상 데이터로 판별되었을 시에는 알림 시스템으로 Timestamp와 해당 데이터의 센

서 명과 이상 데이터를 전달한다.

## 2) 알림 시스템

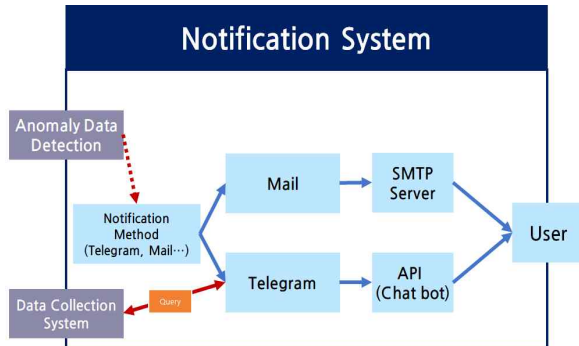


그림 4. 이상 감지 알림 시스템 구조 개요  
 Figure 4. Anomaly Detection Notification System Structure Overview

알림 시스템은 그림 4와 같은 구조를 지니고 있다. 알림 시스템은 Mail 알림과 Telegram 알림 두 가지로 나누어 알림을 보낼 수 있도록 하였다[16]. 메일의 경우 전통적인 알림시스템의 예로써 전달받은 이상 데이터의 Timestamp와 해당 데이터의 센서 명과 이상 데이터를 지정된 메일 주소에 바로 전송할 수 있도록 한다. 메일의 특성상 발송 대기 중인 메일이 많아 과부하가 발생할 경우 문제가 발생할 수 있어 해당 문제의 해결을 위해 발송 제한을 두었다.

Telegram을 활용하는 알림 서비스의 경우 이상이 발생하였을 때는 기존의 이상 데이터 감지 알림 시스템이 호출하는 것과 마찬가지로 호출을 진행하도록 하였다. Mail을 통한 알림 시스템과의 차별점으로 특정 질의를 하면 답변하는 챗봇과 동일한 기능을 구현하였다. 기존의 이상 데이터 감지 시스템에서 알림 시스템을 호출해야 답을 할 수 있었던 알림 시스템이 Telegram 챗봇을 통한 질의 시 데이터 수집 시스템과 직접적인 데이터 교류를 할 수 있도록 하였다. 특정 질의를 챗봇에 보냈을 경우 1분 전에 수집한 데이터의 출력과 같은 작업을 수행할 수 있도록 지원한다.

일부 소스의 수정을 거친다면 여타 널리 사용되는 다양한 메신저, 웹, 어플리케이션 등에도 해당 알림 시스템을 적용할 수 있다.

## 2. 구축 환경

본 연구에서 구축한 이상 데이터 감지 알림 시스템

은 임베디드 디바이스를 활용하고 해당 프로그램을 특정 언어를 사용하여 개발하는 연구를 진행하였다.

### 1) 임베디드 디바이스

본 연구에서 사용한 임베디드 디바이스는 Raspberry Pi 4 model B 메모리 4GB 모델을 사용하였다. 운영체제는 Ubuntu 22.04, Storage는 256GB의 삼성 EVO plus Microsd와 Sandisk 256GB USB를 사용하여 이상 데이터 감지 알림 시스템을 구축하였다.

### 2) 사용 언어

MAIL 서버 및 Telegram API와 연동 및 개발을 하기 위하여 Python을 사용하였으며, 이상 데이터 감지의 통계적 처리를 위해 Python을 활용하여 개발을 진행하였다.

## VI. 결 론

본 연구에서는 데이터 수집 시스템과 데이터 수집 시스템에서 수집한 데이터의 이상을 지속적으로 감시하여 이상 데이터를 확인하여 알림을 주는 시스템을 제안하였다. 실제 실증 중인 공장에서 구동 가능한 수준의 알림 시스템이며, 에너지 관리 시스템이 필수적으로 설치되지 않아도 구동이 가능한 시스템을 제안하여 보편적인 센서 데이터를 사용하는 전 분야에서 활용할 수 있는 시스템을 제안하였다. 특히 공장의 한계를 극복하고자 임베디드 디바이스에서 데이터 수집 시스템과 이상 데이터 감지 알림 시스템 모두를 구축하여 임베디드 디바이스의 활용 한계를 극복하는 데 기여하였다.

향후 연구에서는 현재의 통계학 기반의 알림 시스템의 한계를 극복하고자 인공지능 기법을 활용하여 다양한 데이터에 활용할 수 있도록 비선형적 데이터 패턴 등에 대한 인식을 가능하게 하는 시스템을 임베디드 디바이스에 맞게 경량화하여 개선하여 구축하는 연구를 진행할 예정이다.

## References

[1] "Why industrial complexes are centers of carbon neutrality", SK ecoplant, <https://news.skecoplant.c>

- om/plant-tomorrow/3079/
- [2] Choi, E, Kang, M, Jung, Y, Paik, J. 2017, "Implementation of IoT-based Automatic Inventory Management System", The International Journal of Advanced Culture Technology , vol.5, no.1 pp.70-75. DOI : <https://doi.org/10.17703/IJACT.2017.5.1.70>
- [3] Yong-Soo Lee, Jun Heo, Yong-Hoon Choi "A Study for Space-based Energy Management System to Minimizing Power Consumption in the Big Data Environments", The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 13, No. 6, pp. 229-235, Dec 2013. DOI: <http://dx.doi.org/10.7236/JIIBC.2013.13.6.229>
- [4] Woojin Cho, Hyungah Lee, Chae-young Lim, Dongju Kim, Chae-eun Yeo, & Jae-hoi Gu. (2023). "A Study on the Loading Efficiency of Utility Consumption Prediction Models". Korea Society of Waste Management Spring academic presentations, 2023(0), 515-515.
- [5] Woojin Cho, Chae-eun Yeo, Jae-hoi Gu, Chae-young Lim. (2023). "Evaluation of Storage Engine on Edge-Based Lightweight Platform using Sensor/OPC-UA Simulator". The Journal of the Convergence on Culture Technology (JCCT), 9(3), 803-809. DOI : <https://dx.doi.org/10.17703/JCCT.2023.9.3.803>
- [6] Myeong-Joon Kim, Hyun-Jik Cho, Chul-Goo Kang.(2022). "LSTM-based Anomaly Detection for Screw Air Compressors of Railway Vehicles". Transactions of the Korean Society of Mechanical Engineers - A,46(2),195-202. DOI : <https://dx.doi.org/10.3795/KSME-A.2022.46.2.195>
- [7] Woojin Cho, Chae-young Lim, Jae-hoi Gu. (2023). "Comparison and Evaluation of Data Collection System Database for Edge-Based Lightweight Platform". Journal of Platform Technology, 11(5), 49-58. DOI : <https://dx.doi.org/10.17703/JCCT.2023.9.3.803>
- [8] Se-il Park, Jong-wook Jang.(2017). "Design of smart mobility status notification system". Journal of the Korea Institute of Information and Communication Engineering,21(11),2128-2132. DOI : <https://doi.org/10.6109>
- [9] Kim, Hyeong-Seok, Park, Won-Woo, Lee, Young-Bae, Seo, Na-Hyun, Jeongchang Kim.(2015). "IoT-Based Location Tracking and Emergency Alert System for the Children and the Elderly". The Korean Institute of Broadcast and Media Engineers academic presentations,(0),203-204.
- [10] Youngmee Shin, Tae-Wook Heo, Il-Woo Lee.(2021). "Event Processing Mechanism for Factory Energy Management System(FEMS)". Proceedings of Symposium of the Korean Institute of Communications and Information Sciences,(0),297-298.
- [11] Unified Architecture, "OPC Foundation", <https://opcfoundation.org/about/opc-technologies/opc-ua/>
- [12] MySQL, "InnoDB Storage Engine", <https://dev.mysql.com/doc/refman/8.0/en/innodb-storage-engine.html>
- [13] RocksDB, "RocksDB", <https://rocksdb.org/>
- [14] Raspberry PI 4 model B, "Raspberry PI 4 model B", <https://www.raspberrypi.com/products/raspberry-pi-4-model-b/?variant=raspberry-pi-4-model-b-4gb>
- [15] MyRocks, "Facebook", <https://github.com/facebook/mysql-5.6>
- [16] Telegram API, "Telegram", <https://core.telegram.org/>

※ 본 연구는 산업통상자원부(MOTIE)의 재원으로 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구입니다. (No.2020202900170)