

A Study on the Build of Equipment Predictive Maintenance Solutions Based on On-device Edge Computer

Yong-Hwan Lee*, Jin-Hyung Suh**

*CEO, IANSIT Co., Ltd., Incheon, Korea

**Professor, Dept. of Tax Accounting, Kyungin Women's University, Incheon, Korea

[Abstract]

In this paper we propose an uses on-device-based edge computing technology and big data analysis methods through the use of on-device-based edge computing technology and analysis of big data, which are distributed computing paradigms that introduce computations and storage devices where necessary to solve problems such as transmission delays that occur when data is transmitted to central centers and processed in current general smart factories. However, even if edge computing-based technology is applied in practice, the increase in devices on the network edge will result in large amounts of data being transferred to the data center, resulting in the network band reaching its limits, which, despite the improvement of network technology, does not guarantee acceptable transfer speeds and response times, which are critical requirements for many applications. It provides the basis for developing into an AI-based facility prediction conservation analysis tool that can apply deep learning suitable for big data in the future by supporting intelligent facility management that can support productivity growth through research that can be applied to the field of facility preservation and smart factory industry with integrated hardware technology that can accommodate these requirements and factory management and control technology.

▶ **Key words:** Smart Factory, Edge Computing, On-Device, Distributed Computing, Equipment Predictive Maintenance, Productivity

[요 약]

본 연구에서는 현재 일반적인 스마트 팩토리에서 데이터 전송에 사용하는 중앙 집중형 시스템에서 발생하는 데이터를 중앙의 센터까지 전송, 처리할 때 발생하는 전송 지연 등의 문제 해결을 위하여 필요한 곳에 연산과 저장 장치를 도입하는 분산 컴퓨팅 패러다임 (Distributed Computing Paradigm)인 온-디바이스 (On-Device) 기반 에지 컴퓨팅 (Edge Computing) 기술과 빅데이터 분석 기술 및 활용 방법의 연구를 통하여 설비 고장 등을 예지하여 가동율을 높일 수 있는 산업현장의 설비관리에 활용되는 솔루션을 제안한다. 그러나 에지 컴퓨팅 기반의 기술이 실제 적용되더라도 네트워크 에지에서 장치의 증가는 많은 양의 데이터가 데이터 센터로 전달되어 네트워크 대역이 한계치에 이르게 되어 네트워크 기술의 향상에도 데이터 센터는 수많은 응용에서 중요한 요건이 되는 수용 가능한 전송 속도와 응답 시간을 보장하지 못하게 된다. 이와 같은 요구조건을 수용할 수 있는 일체형 하드웨어 기술과 공장관리 및 제어 기술을 적용한 설비보존 및 스마트 팩토리 산업 분야에 적용할 수 있는 연구를 통하여 생산성 증대를 지원할 수 있는 지능적 설비관리를 지원하도록 하여 추후 빅데이터에 적합한 딥러닝을 적용할 수 있는 인공지능 기반 설비 예지 보전 분석 도구로 발전할 수 있는 기반을 제공한다.

▶ **주제어:** 스마트 팩토리, 에지 컴퓨팅, 온-디바이스, 분산 컴퓨팅, 설비 예지 보전, 생산성

- First Author: Yong-Hwan Lee, Corresponding Author: Jin-Hyung Suh
- *Yong-Hwan Lee (kaeby@iansit.co.kr), IANSIT Co., Ltd.
- **Jin-Hyung Suh (jhsuh@kiwu.ac.kr), Dept. of Tax Accounting, Kyungin Women's University
- Received: 2020. 03. 24, Revised: 2020. 04. 14, Accepted: 2020. 04. 16.

I. Introduction

스마트 팩토리 (Smart Factory)는 전통 제조산업에 정보통신기술을 결합하여 개별 공장의 설비 (장비) 공정이 생산 네트워크로 연결되고, 모든 생산 데이터와 정보가 실시간으로 공유 및 활용되어 최적화된 생산 운영이 가능한 공장으로, 공장 간의 협업 운영이 지속되는 생산 체계를 의미하며, [1] 이에 산업현장에서 공정 및 품질관리를 위한 다양한 센서를 설치하고 이를 통해 효과적 생산활동을 수행하고 있다.

최근에는 센싱 기술의 발전을 통해 기존에는 획득할 수 없었던 다양한 종류의 데이터에 대한 실시간 수집이 가능해짐에 따라 실시간 빅데이터를 효율적으로 관리하고 의미 있는 결과를 도출하기 위한 필요성이 높아지고 있으나, [2] 빅데이터 분석을 위해서는 데이터를 저장하고 관리하기 위한 인프라 구축에 많은 시간과 비용이 소요되고 분석 전문가에 관한 기술 의존도가 높은 데이터 분석 분야의 특성상 여러 산업현장에서 빅데이터 기반 분석 기술을 활용하는 데 어려움이 크다. [3]

스마트 팩토리 분야에서 가장 큰 문제점은 각 산업 분야별 생산 공정을 이해하고 빅데이터 기술을 어떻게 접목하여 기업의 문제를 해결할지 방안을 제시하기 어렵다는 데 있다. 이는 데이터 분석 전문가가 해당 분야에 대한 전문 지식을 갖추기 어렵고, 반대로 산업현장의 전문가는 데이터 분석에 대한 이해가 높지 않기 때문이다.

이러한 문제를 해결하기 위해 데이터 분석 전문가 외에도 전문 지식이 없는 비전문가도 활용할 수 있는 공정 설비 기반 빅데이터 분석 플랫폼이 필요로 하며, 필요로 하는 플랫폼은 기본적으로 생산활동의 시작과 끝을 의미하는 물리적 또는 논리적인 공간정보 (설비) 기반으로 센싱 데이터를 수집, 저장, 관리하고 공정 설비에 위치하는 센서의 물리적인 순서 또는 논리적인 순서를 고려한 분석을 수행할 수 있도록 지원하는 것을 목적으로 한다. 그리고 플랫폼을 통해 전체 생산활동 (공정 설비)을 하나의 주기로 보았을 경우 어느 주기 (시점)에서 품질문제를 발생시키는지 분석하여 개선할 수 있도록 지원하여야 한다.

II. Preliminaries

1. Related works

1.1 Smart Factory

전통적으로 스마트 (Smart)라는 단어는 감지 (Sensing), 통제 및 작동 (Control and Actuating) 기능

을 가진 객체에 부여되었던 수식어이지만 현재는 그 의미가 퇴색되었다고 할 정도로 남발하고 있는 것이 사실로 실상 제조업에서의 스마트화는 아직 구체화된 실체가 부족한 상황에서 어휘를 선점하려는 경향을 보이고 있다. 이러한 상황에서의 스마트 팩토리에 대한 정의는 각 기관과 주제마다 차이가 있으며, 스마트 팩토리가 지니는 사전적 의미는 “영리한 + 공장”이라 할 수 있으나, 이를 다시 풀어 보면 생산 전략에 기반을 둔 제조 여건 변화에 유연하게 대응하고, 공급망 관리 (Supply Chain Management, SCM) 통합 관점의 QCD (Quality, Cost, Delivery) 및 제약 관리로 생산 운영을 신뢰성 있게 수행하는 공장이라고 할 수 있다. 스마트 팩토리의 정의를 그 특징과 요건을 통해 살펴보면 다음과 같다. [4]

(1) 신뢰성: 수집된 데이터의 신뢰성에서 시작해 인터록 (Interlock) 등 작업 이상 상황에 대한 안정성, 예측 가능한 작업 수행 보장까지의 역할 수행한다.

(2) 능동성: 신규 데이터 상관성 도출, 재고 감축 작업 지시, 장기 재고 이적 등 판단 결과에 기반을 둔 이행 기능 수행한다.

(3) 지능성: 변화된 여건에 따라 스스로 판단하는 의사결정력을 발휘하며 일괄 처리작업 (Batch Job)을 통한 주기적 판단과 마스터 판단 기준을 통한 입력 데이터에 따른 출력을 도출한다.

(4) 연계성: 다양한 대량의 데이터를 유관 데이터 영역으로 검토해 활용하며, 기능과의 연계로 (예: 수집, 저장, 가공, 활용 사이클상의 복잡한 이벤트 처리 (CEP: Complex Event Processing) 데이터 영역을 확대 운영한다.

(5) 민첩성: 실시간 처리 수준의 향상과 제조 운영 관련 경보, 조치 소요시간, 정보 공유 등의 기능의 빠른 대응력을 확보한다.

1.2 Big Data Platform

빅데이터 플랫폼은 빅데이터 기술의 집합체이자 데이터 기술을 잘 사용할 수 있도록 준비된 환경으로 빅데이터 세부 기술의 집합체로 볼 수 있으며 빅데이터를 수집하고 저장하고, 처리하고 관리하는 일련의 과정을 통합적으로 제공해야 하며, 산업현장에서 발생하는 빅데이터의 경우 단위 공장 또는 특정 조직의 목적에 맞게 별도로 수집하기보다는 전사의 데이터를 한곳에 모으기 위해 이와 같은 빅데이터 플랫폼이 필요하다. [5]

스마트 팩토리에서 적용할 수 있는 빅데이터 플랫폼 기술은 크게 하드웨어와 소프트웨어 기술로 구분할 수 있다. 먼저 하드웨어 환경은 빅데이터 수용 용량 및 처리, 분석 작업에 대한 부하 등을 고려하여 설계되어야 하는데 컴퓨

팅 부하 처리를 효과적으로 수행하는 빅데이터 플랫폼은 빅데이터의 분석 결과의 활용 목적에 따른 분석 방법과 기술 및 플랫폼 구성은 용도에 따라 다양하게 설계할 수 있다. 다음으로 빅데이터 플랫폼의 소프트웨어 기술은 빅데이터의 분석에 필요한 수집, 관리 분석, 사용자 환경 등에 대한 소프트웨어를 구축하게 되는 것으로 데이터의 처리 및 분석 모듈, 데이터의 수집 및 정제 모듈, 서비스 관리 모듈, 사용자 관리 모듈, 모니터링 모듈, 보안 모듈 등으로 구성되며, 빅데이터 처리 및 분석과 이를 위한 데이터의 수집 및 정제 등을 수행한다.

일반적으로 아파치 하둡 (Apache Hadoop) [6][7]과 같은 오픈 소스 기반의 빅데이터 시스템에서 자료수집에서부터 저장, 처리, 분석 등 각 영역에서 사용하는 에코시스템 (Echo System)을 도입하여 사용되며 선택에서 적용, 최적화 및 운영에 이르기까지 모든 과정을 사용자 스스로가 수행하고 책임져야 한다. 이러한 관점으로 구성할 수 있는 기업 내의 빅데이터 플랫폼은 다음의 Fig. 1과 같다 [8].

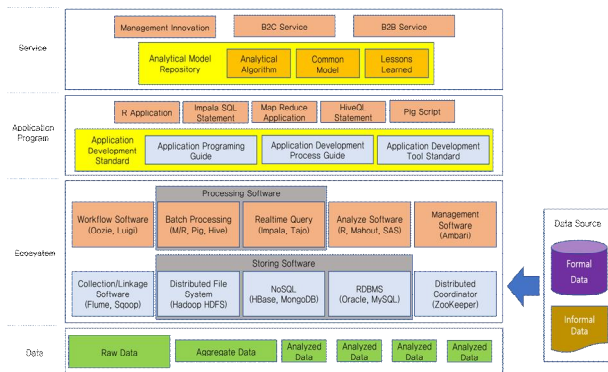


Fig. 1. Big Data Platform

III. The Proposed Scheme

기본적으로 스마트 팩토리 환경에서 제공할 수 있는 빅데이터 통합 플랫폼은 크게 실시간 센싱 빅데이터 수집 (I/O Server), 솔루션을 다양한 현장에서 사용할 수 있도록 지원하는 저작도구 (Smart Factory Editor), 빅데이터 분석을 통한 설비 에지 컴퓨팅 기술 (Edge Computing), 빅데이터 통합 관리기술 (Big Data Integration Management) 및 빅데이터 기반 가시화 및 서비스 기술로 구성되며, 각 기술에 대한 특징은 다음과 같다.

3.1 Real-time sensing big data collection

스마트 팩토리에 있어 빅데이터 처리는 비정형 데이터 수집에 있어 병목 현상이 발생하여, 장기간의 데이터 저장

에서의 인프라 한계 및 전수 데이터의 분석 요구라는 문제점을 해결하여야 한다. 이러한 문제점을 해결하기 위하여 실시간 센싱 데이터 수집을 수행하게 되며, 이를 통하여 수평 확장 (Scale-out) 구조의 실시간 수집 기능과 빅데이터 플랫폼으로의 장기 데이터 저장 기반을 마련할 수 있고 전수 데이터의 고급 분석을 통하여 신뢰성을 향상할 수 있다.

3.1.1. Collection of previous sensing data

제조현장에서 발생하는 RFID, 바코드, GPS, 온/습도, 압력, CO2 등 물리/화학적 정보와 같은 다양한 종류와 아날로그 성격의 데이터는 다양한 설비 시스템 제조사에 의해 통신 프로토콜이 혼재되어 있고 대량생산체계에 따른 데이터 발생 속도가 빠르고 대용량이라는 특징이 있다. 특히 공정별 설비환경의 차이로 인한 자료 형식의 다양성이 존재하여 다양한 자료를 처리할 수 있는 개방형 인터페이스 기능이 필요하다.

지금까지 현실 세계의 물리적 상황 정보를 수집하기 위한 개방형 인터페이스의 연구로는 RFID/USN 및 사물 통신 (M2M; Machine-to-Machine) [8] 분야를 중심으로 이루어져 왔는데 RFID/USN에서는 사물의 이력 정보뿐만 아니라 사물을 둘러싸고 변화하는 물리 환경 계의 다양한 정보를 획득하여 생산성, 안전성 및 인간 생활 수준의 고도화를 실현한다. [9] 사물 통신은 단말로부터 수집된 정보를 무선 접속 네트워크와 이동통신 네트워크를 통한 정보 전달 서비스로 적용됐으나 사물 통신 장비에서 수집한 정보를 가공하여 사용자에게 제공하는 수동적인 개념의 서비스 형태가 주로 이루어졌던 한계점이 있다.

3.1.2. IoT-based sensing information collection

USN과 사물 통신의 수동적인 형태의 서비스를 통한 다양한 자료의 형식을 제대로 처리할 수 없어 데이터의 수집과 통신능력을 이용하여 물리적, 가상적 대상들을 연결하는 글로벌 네트워크 인프라인 개방형 사물인터넷 (IoT, Internet of Things) 플랫폼을 이용하게 되는데 이 사물인터넷 응용을 개발하는 데 있어 실시간 통신 기술은 필수적인 요소다. 실시간 커뮤니케이션의 인스턴트 메시지 기능을 지원하기 위한 IoT 표준 프로토콜에는 CoAP 및 MQTT 프로토콜을 사용할 수 있다.

- CoAP (Constrained Application Protocol) [10]

IoT의 중요성이 증가하면서 IETF는 경량 메시지전달을 위한 CoAP를 정의하였는데 IETF의 정의에 따르면, CoAP는 '제한된 노드와 제한된 (즉, 저전력의 손실이 많은) 네트워크'를 위한 프로토콜이다.

워크'를 위한 것으로 이클립스 (Eclipse) 커뮤니티에서도 MQTT와 마찬가지로 CoAP를 공개 표준으로 지원하고 있다. CoAP는 상용 적으로 지원되며, IoT 제공업체들과 함께 빠르게 성장하고 있다. CoAP는 클라이언트/서버 프로토콜이며, 일대일 (1:1) '요청/보고' 상호작용하는 모델을 제공한다. 또한, 아직 IETF 표준화의 초기 단계에 있지만, 멀티캐스트를 지원한다. 십여 년 전 개발된 프로토콜로부터 IoT의 요구에 맞추어 개조된 MQTT와 달리 CoAP는 IETF가 처음부터 제한된 환경에서 동작하는 제한된 기기의 경량 메시지처리용 IoT를 지원하기 위해 만들어졌다. CoAP는 간단한 프록시를 통해 HTTP와 RESTful 웹과 상호 운용되도록 설계되어 있어 본질에서 인터넷에 적합하다.

● MQTT (Message Queuing Telemetry Transport)

제한된 네트워크 및 다중 플랫폼 환경에서 성능이 제한된 노드들이 경량의 메시지로 통신할 수 있도록 설계된 프로토콜이다. 다른 프로토콜에 비해 작은 고정 헤더와 오버헤드의 최소화로 제한된 환경에서 사용하기 적합한 프로토콜로 [11] 기본적으로 신뢰성과 저전력을 특징으로 하므로 IoT 네트워크에 적용하기 적합하다.

통합 플랫폼에서 처리하는 센싱 빅데이터는 스마트 팩토리를 대상으로 하며 설비 내 센싱 정보로 구성되어 있다. 센싱 정보는 설비에 직접 설치된 센서로부터 측정된 온도, 소음 등의 정보이며, 이를 실시간으로 수집되는데 수집되는 센싱 정보는 유선 또는 무선을 이용하여 수집하여 MQTT로 전달되는데 전달된 센싱 정보를 효과적으로 수집하기 위해 아파치 플럼-NG (Apache Flume-NG)를 사용한다. 기본적으로 아파치 플럼은 에이전트 계층 (Agent Layer), 컬렉터 계층 (Collector Layer), 스토리지 계층 (Storage Layer)의 총 3계층으로 구성되어 있는데 에이전트 계층에서 각 에이전트 노드는 수집할 로그 데이터 (Log Data)가 생성되는 기계에 설치하는 것이 일반적이다. 대상 기계가 여러 개일 경우 기계마다 에이전트 노드 (Agent Node)를 설치하고 이 에이전트 노드가 에이전트 계층을 형성한다. 에이전트 노드에서 수집한 데이터는 컬렉터 노드로 전송되는데 컬렉터 노드는 여러 컬렉터 노드로 구성할 수 있으며, 에이전트 노드에서 컬렉터 노드로 데이터를 전송할 때는 어떤 데이터를 어디로 보내고 어떻게 처리할 것인지 등에 대한 데이터 흐름 (Data Flow)를 설정할 수 있고 이 설정대로 데이터를 이동시켜 스토리지 계층에 저장한다. 수집된 센싱 정보가 일정한 크기에 도달하거나 특정한 시간별로 데이터를 관리할 필요가 있는 경우에는 기존 데이터의 쓰기를 중지하고 새로운 공간에 데이터를 저장하여야 한다. 본 연구에서 센싱 정보를 수집하기 위해 적용한 플럼 에이전트 속성은 Table 2와 같다.

Table 1. Analysis of CoAP and MQTT

Item	CoAP	MQTT
Communication Method	Request/Response	Publish/Subscribe
Communication Structure	Server-Client	Publisher-Broker-Subscriber
Encryption Method	DTLS	TLS
E2E Encryption Support	Support	Impossible
Message Reliability (QoS)	Support	Impossible
Group Communication	Possible	Possible
Contents Type	Various (XML, JSON, TLV, plain-text, & etc)	XML
Transmission Protocol	UDP	TCP
Application Level Standards	Resource Model (RESTful)	Unsupported

3.2 Authoring tool

솔루션을 다양한 현장에서 사용할 수 있도록 지원하는 저작도구는 데이터 과학 (Data Science)를 위한 것으로 수집된 빅데이터의 크기가 일반적으로는 분석할 수 없을 정도로 커졌고, 분석할 수 없을 정도로 데이터가 복잡해짐으로 인하여 이론적 사고, 실험적 사고, 컴퓨팅 사고를 체계적으로 교육받지 못한 사용자에서 빅데이터를 통한 가치를 최대한 많이 즉시 창출할 수 있도록 한다.

본 연구에서 제안된 사용자 중심의 설비 등록 기능과 IoT 기반 계측 센서, 설비 등록 및 데이터 구조 설정 기능과 설비 등록 도구 제작에 따른 상세 설계 및 데이터베이스의 사용 그리고 사용자 중심의 UI/UX 개발로, 화면상으로 장비 등록과 설정만으로 손쉽게 외부 시스템과 연결할 수 있도록 구현된다.

3.3 Equipment predictive maintenance technology

빅데이터 분석 기술을 통하여 설비의 상태를 확인하거나 결함을 분석하고 설비 상호 간의 연관성을 파악하게 되고 여기에 이러한 분석 데이터들을 전문가의 지식 또는 지식이 적용된 자동 진단 알고리즘에 접목함으로써 설비의 결함을 자동으로 진단하고 결함이 발생한 원인과 진전 과정, 운전상의 문제점 및 위험 정도 파악 등 전문가가 할 수 있는 의사결정을 대신하여 현장 관리자에게 알려주거나 전문가가 정확한 판단을 할 수 있도록 지원한다.

3.3.1 Big data integrated management technology

실시간으로 수집되는 빅데이터를 효과적이고 통합적으로 관리하여 가치 있는 데이터를 지속해서 발굴하는 기술을 의미하며 데이터 관리 소홀로 인한 정보 노출과 관리 비용을 줄이고, 고품질의 데이터를 발굴하여 기업의 가치 창출에 이바지하기 위해서는 전사적으로 데이터를 관리하고 통제한다.

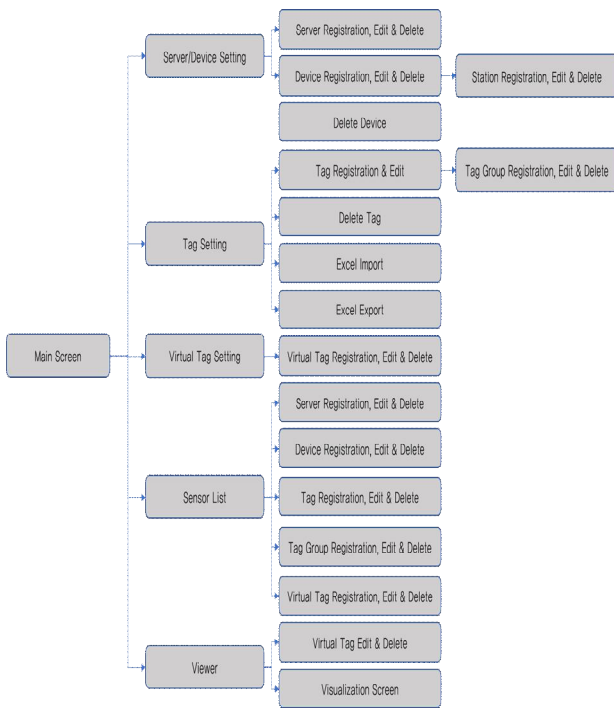


Fig. 2. Interface structure

3.3.2 Big data-based visualization and service technology

가시화 및 그 서비스 기술은 데이터 분석 결과를 더욱 쉽게 이해할 수 있도록 그래프, 도표 등으로 사용자가 더욱 쉽게 데이터의 정보를 파악할 수 있어, 즉각적인 판단이 가능하게 하여 빠른 의사결정을 할 수 있도록 한다. [9] 즉, 장비로부터 전송되는 데이터를 시스템에 맞는 데이터로 설정했다면 해당 장비로부터 전송되는 데이터를 가시화해주는 단계로 전송을 받은 데이터를 활용하여 데이터를 전송한 장비가 위치한 곳에 이상 유무를 판단해 이상이 있는지 없는지에 대해 가시화한다.

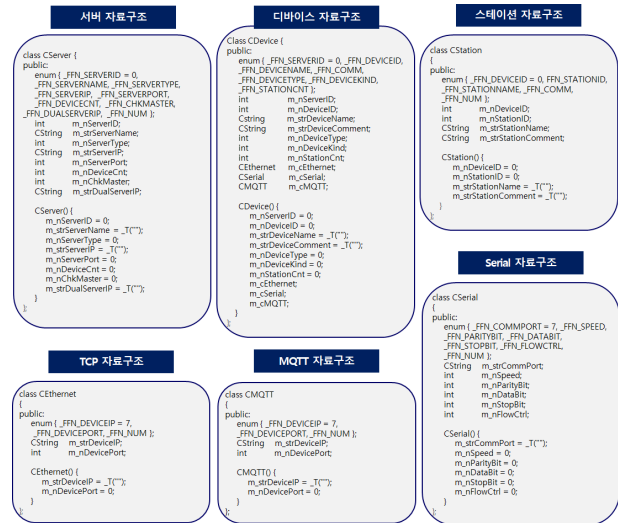


Fig. 3. Data structure

IV. Performance

본 연구는 빅데이터 기반의 설비보전 예측 플랫폼을 위한 것으로 센싱 빅데이터의 수집, 저장, 분석 및 시각화 기능을 갖춘 플랫폼 구축 개발을 목적으로 하고 있으므로 이 연구에 대한 성능 테스트는 다량의 다종 데이터가 입력되는 빅데이터의 특성을 충족시키는지를 확인하였다.

4.1 Equipment sensing information collection test

가상 설비 센싱 데이터를 개발된 빅데이터 시스템을 통해 HDFS (Hadoop Distributed File System)에 저장 시 21.1GByte 크기의 가상 설비 센싱 데이터를 HDFS에 저장하는데 약 35분의 시간이 소요되고 빅데이터의 경우 조인 연산을 수행하여 저장하는 것보다 원본 그대로 저장하는 방식이 더 적합함을 확인할 수 있다.

4.2 Real-time collection limit testing

수집 인터페이스 기능이 설비 센싱 데이터를 실시간으로 어느 정도 수집이 가능한가를 확인하기 위하여 가상으로 설비 센싱 데이터 생성기를 PC로 구축하였으며 시험 환경 개념도는 Fig. 6과 같으며 PC 1대당 각각 2만 건씩 총 4만 건의 설비 센싱 데이터를 빅데이터 마스터 서버로 TCP 방식을 통해 전송했고 시험결과 서버 1대 (포트 1개) 당 분당 최대 4만 건의 데이터를 수집할 수 있었으나 서버 1대를 통해 분당 5만 건의 정보를 수집한 경우 수집에 실패했다. 따라서 가변적인 서버 환경을 구축한다면 보다 대규모 정보 수집이 가능할 것으로 확인되었다.

Table 2. Attributes of Flume Agent

```

AgtRtCar.sources = RtCar
AgtRtCar.channels = CassandraChannel HDFSCannel
AgtRtCar.sinks = Cassandra HDFS
AgtRtCar.sources.RtCar.type =
com.hkinet.flume.source.RtCarSource
AgtRtCar.sources.RtCar.channels = CassandraChannel
HDFSCannel
AgtRtCar.sources.RtCar.selector.type = replicating
AgtRtCar.sinks.Cassandra.type =
com.hkinet.flume.sink.CassandraSinkRtCar
AgtRtCar.sinks.Cassandra.channel = CassandraChannel
AgtRtCar.sinks.HDFS.channel = HDFSCannel
AgtRtCar.sinks.HDFS.type = hdfs
AgtRtCar.sinks.HDFS.hdfs.path =
hdfs://master.hdp.hkinet.com:8020/user/kict/openapi/rtc
ar/%Y-%m-%d
AgtRtCar.sinks.HDFS.hdfs.filePrefix = RealTime_Car
AgtRtCar.sinks.HDFS.hdfs.fileType = DataStream
AgtRtCar.sinks.HDFS.hdfs.writeFormat = Text
AgtRtCar.sinks.HDFS.hdfs.batchSize= 10000
AgtRtCar.sinks.HDFS.hdfs.rollSize = 10485760
AgtRtCar.sinks.HDFS.hdfs.rollCount = 0
AgtRtCar.sinks.HDFS.hdfs.rollInterval = 600
AgtRtCar.sinks.HDFS.hdfs.minBlockReplicas = 1
AgtRtCar.channels.CassandraChannel.type = memory
AgtRtCar.channels.CassandraChannel.capacity = 10000
AgtRtCar.channels.CassandraChannel.transactionCapacity =100
AgtRtCar.channels.HDFSCannel.type = memory
AgtRtCar.channels.HDFSCannel.capacity = 10000
AgtRtCar.channels.HDFSCannel.transactionCapacity = 100
    
```

```

Map-Reduce Framework
Map input records=171226646
Map output records=171226646
Input split bytes=87
Spilled Records=0
Failed Shuffles=0
Merged Map outputs=0
GC time elapsed (ms)=8684
CPU time spent (ms)=1038980
Physical memory (bytes) snapshot=452931584
Virtual memory (bytes) snapshot=4410298368
Total committed heap usage (bytes)=219152384
File Input Format Counters
Bytes Read=0
File Output Format Counters
Bytes Written=2747895703
17/06/24 16:04:29 INFO mapreduce.ImportJobBase: Transferred 21.1856 GB in 2,093.1797 seconds
17/06/24 16:04:29 INFO mapreduce.ImportJobBase: Retrieved 171226646 records.
    
```

Fig. 4. Store the virtual facility sensing data

4.3 Equipment maintenance prediction test through noise

설비보전 예지 테스트를 위해 산업현장 컨베이어벨트의 메인 구동축의 모터로부터 소음 정보를 수집하였고 이를 양자화 후 0.2초 간격으로 표본화하였다. 모터의 작동 상태는 정지 상태에서 컨베이어벨트를 정상 운용 가능한 상태까지 시작하는 단계, 정상 운용 가능한 상태 단계, 1차 이상 상태 단계 (간섭으로 인한 정상 운용 상태보다 불규칙한 소음 발생), 2차 이상 상태 단계 (간섭으로 인한 1차 이상 상태 단계보다 불규칙한 큰 소음 발생) 및 정상 운용 가능 상태에서 정지 상태로 종료하는 단계로 구분할 수 있으며 다음의 표에서와같이 각 단계는 50개의 데이터와 단계별 표본화는 연속된 10개, 각 양자화된 수치의 범위는 0~125 사이로 생성되었으며 정상 데이터는 1, 2, 5단계, 비정상 데이터는 3, 4단계이다.

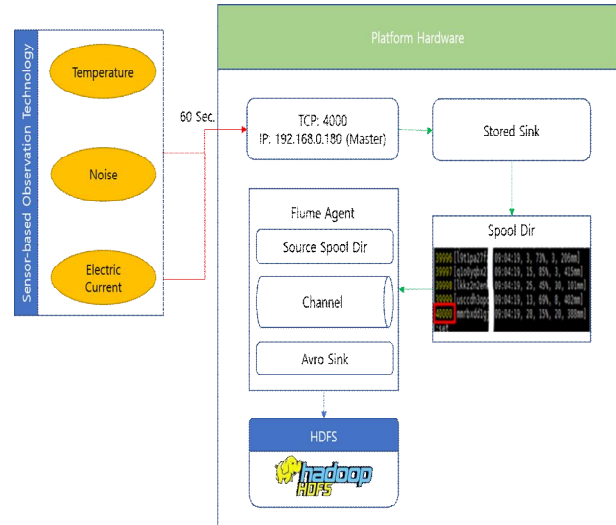


Fig. 5. Virtual sensing data generation environment

수집된 데이터를 정상/비정상 상태를 인식하기 위해 각 5단계를 구분해야 하는데 이를 위해 기계학습 기반의 알고리즘인 KNN (K-Nearest Neighbor), LR (Linear Regression), SVM (Support Vector Machine), GNB (Gaussian Naive Bayes) 4개 알고리즘을 통해 학습 및 분류하였다.

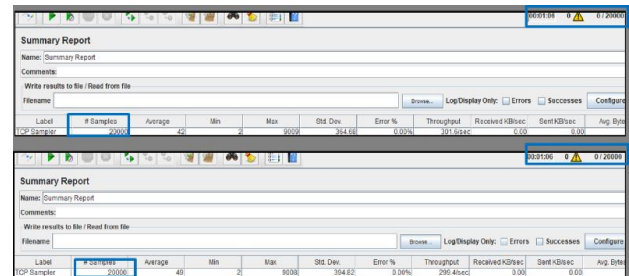


Fig. 6. Real-time collection of virtual sensing data

Table 3. Step-by-step collection data

Step	1	2	3	4	5	6	7	8	9	10
Step 1 - 1	4	14	28	36	36	51	55	67	74	76
Step 1 - 2	12	17	21	33	37	45	59	61	76	81
Step 2 - 1	81	77	79	79	80	82	80	84	82	78
Step 2 - 2	78	77	80	84	80	84	78	81	82	80
Step 3 - 1	92	76	88	79	71	94	90	63	75	62
Step 4 - 2	68	92	84	87	79	91	87	69	83	87
Step 5 - 1	59	54	70	52	93	42	51	83	70	70
Step 5 - 2	89	74	93	76	88	121	98	114	120	42

총 250개 (단계별 50개) 데이터를 단계별로 학습용 80% (200개), 테스트용 20% (50개) 로 구분하였고 데이터 별 입력은 10개이며, 출력은 1개로 출력값은 1~5단계를 나타낸다. 단, 입력 데이터의 특징을 보다 뚜렷하게 학습하기 위해 10개의 입력 데이터를 통계함수를 이용해 평균, 표준편차, 최솟값, 25%, 50%, 75%, 최댓값을 7개의 입력 데이터로 바꾸어 처리하도록 설계하였다.

V. Conclusions

본 연구에서는 스마트 팩토리 시스템에 기반한 설비 예지 보전을 구축하기 위해 설비 센싱 빅데이터 처리를 위한 플랫폼을 개발하고 이를 효과적으로 가시화하기 위한 기술을 개발하였다. 먼저, 실시간으로 생성되는 센싱 정보인 온도, 소음, 전류 등을 수집하기 위해 MQTT 기반 실시간으로 수집 가능한 기술과 풀럼 기반 센싱 정보 수집 모듈을 개발하였으며, 대량의 데이터 수신과 저장을 할 수 있게 했다. 또한, 빅데이터를 신속하게 분석할 수 있도록 기반을 구축했다.

의 적용과 개발된 분석 기능 중 추정치와 같은 분석 기능이 통합 플랫폼에 활용될 수 있음을 보여주었다.

시범 구축한 플랫폼의 성능시험을 위해 수집한 빅데이터를 사용하여 수집된 정보를 가공하여 가상의 대용량 센싱 정보로 변환하였고, 생성된 빅데이터를 실시간 수집하고 저장, 분석하는 시험을 수행하였다. 마지막으로 설비 예지 보전을 위해, 소음 데이터를 기계학습 기반의 알고리즘을 통해 확인해 본 결과 1~5단계까지의 상태를 최대 82%의 정확도로 찾아낼 수 있었다. 이를 활용해 이상 상태로 감지할 수 있는 3, 4단계를 사전에 알려줌으로써 설비 예지 보전이 가능할 수 있음을 확인할 수 있었다.

향후 현재 시험대에서 수집 중인 설비 센싱 데이터가 전국적으로 확장되었을 경우를 대비하여 네트워크 균형을 고려한 센싱 정보의 실시간 수집 환경을 구축과 업그레이드를 통해 처리속도 및 발열 등의 문제점을 개선할 필요가 있다. 또한, 무선 통신망을 통한 정보 수집에 대한 IoT 표준 구조 기반의 CoAP 등과 같은 다양한 프로토콜을 지원에 대한 사업 개발이 필요하며, 표준기반의 무선 프로토콜 지원을 통해 이동기기 분야 이외 다양한 분야에 사업 성과를 적용 가능성을 확인하는 사업이 요구된다. 또한, 통합 플랫폼 내의 분석 도구는 분석과 예측의 정확도를 높이기 위해 딥러닝 과정이 필요하며, 이를 위해 향후 기계학습 시스템을 조사 후 빅데이터에 적합한 딥러닝을 적용할 수 있는 인공지능 기반 설비 예지 보전 분석 도구로 발전할 필요가 있다.

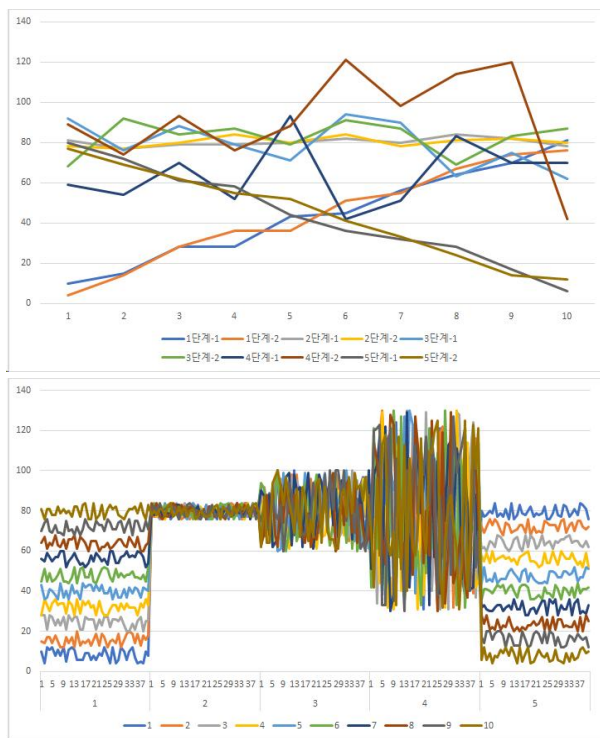


Fig. 7. Visualization for analysis of collected data

실시간으로 생성되는 센싱 빅데이터를 융합, 활용할 수 있는 설비 예지 보전을 분석하고 이를 가시화할 수 있는 기능 및 개발한 통합 플랫폼을 통해 다양한 분석 기능의 적용이 가능한지 실증하기 위해 설비보전 추정 분석 기능

Table 4. Algorithm-specific accuracy

Alg.	Step	Accuracy	Reproducibility	F1-Score
KNN	1	0.60	0.30	0.40
	2	1.00	1.00	1.00
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	0.53	0.80	0.64
Accuracy: 0.82				
SVM	1	0.29	0.20	0.24
	2	1.00	1.00	1.00
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	0.38	0.50	0.43
Accuracy: 0.74				
LR	1	0.00	0.00	0.00
	2	0.40	1.00	0.57
	3	0.10	0.20	0.13
	4	1.00	0.40	0.57
	5	0.00	0.00	0.00
Accuracy: 0.32				
GNB	1	0.25	0.20	0.22
	2	1.00	1.00	1.00
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	0.33	0.40	0.36
Accuracy: 0.72				

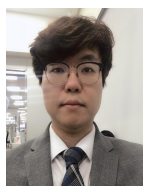
ACKNOWLEDGEMENT

This work was supported by the Technology development Program (S2610530) funded by the Ministry of SMEs and Startups (MSS, Korea)

REFERENCES

- [1] Kyu-Teak Lee, "Smart Factory Industrial R & D Strategy", Ministry of Trade, Industry and Energy, <http://www.krnet.or.kr/board/data/dprogram/1924/B21-%C0%CC%B1%D4%C5%C3.pdf>, 6. 2015
- [2] Tao, F., Cheng, Y., Da Xu, L., Zhang, L., and Li, B. H., "CCIoT-CMfg: Cloud Computing and Internet of Things-Based Cloud Manufacturing Service System", IEEE Transactions on Industrial Informatics, 10(2), pp.1435-1442, Feb. 2014. DOI: 10.1109/TII.2014.2306383
- [3] Avita Katal, Mohammad Wazid and R. H. Gouda, "Big data: Issues, challenges, tools and Good practices", 2013 Sixth International Conference on Contemporary Computing (IC3), Aug. 2013. DOI: 10.1109/IC3.2013.6612229
- [4] Seung Taek Kim, "Considerations for the successful introduction of smart factories", Deloitte Korea Review, https://www2.deloitte.com/content/dam/Deloitte/kr/Documents/insights/deloitte-anjin-review/07/kr_insights_deloitte-anjin-review-07_05.pdf, pp.37-45, 2016.
- [5] Won-Joong Jang, Sung-In Cho, Soo-Sang Kim, Gwang-Yong Gim, "A Study on the Implementation of Big Data Infrastructure in Smart Factory", Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology Vol.8, No.10, pp.11-23, Oct. 2018
- [6] Apache Hadoop 3.2.1, <https://hadoop.apache.org/docs/current/>
- [7] C.L. Philip Chen, Chun-Yang Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data", Information Sciences 275, pp314-347, 10 Aug. 2014,
- [8] "What is Big Data Platform?", Biz & Tech ICT Report, <https://blog.skcc.com/1734>, 4. 2014.
- [9] Sun Jin Kim, Seok Ji Park, Nae Su Kim., "A Trend Analysis of RFID/USN Industry", Electronics and Telecommunications Trends Vol.20 No.3, pp.43-55, 6. 2005
- [10] "MQTT: the open road to internet of things", <https://www.ekito.fr/people/mqtt-the-open-road-to-internet-of-things/>
- [11] Sun Hee Park, Jeong Ho Kim, Hyun Bae Ryu, "Implementation of public data contents using Big data Visualization technology - Map visualization technique", Journal of Digital Contents Society v.18 No.7, pp.1427 - 1434 30 Nov.. 2017.
- [12] Jun-Sung Park, Dong-Gie Kim, "New Generation Mobile Service: M2M Services and Technology," Comm. of KIISE, Vol.28, No.9, pp.28-39, Sep. 2010.
- [13] Koojana Kuladinithi, Olaf Bergmann, Thomas Pötsch, Markus Becker, Carmelita Görg, "Implementation of CoAP and its Application in Transport Logistics", ReearchGate, <https://www.researchgate.net/publication/229057545>, Jan. 2011

Authors



Yong-Hwan Lee received the B.S. and M.S. degrees in Computer Science from Inha University, Korea, in 1996 and 1998, respectively. Ms. Lee worked at KGI Co., Ltd. Korea, from 1999 to 2016.. He is

currently a CEO in IANSIT Co., Ltd. Korea. He is interested in building information modeling geographical information system, and artificial intelligence.



Jin-Hyung Suh received the B.S and, M.S. degrees in Computer Science from Hongik University, Korea, in 1986 and 1988 respectively and Ph.D degrees in Computer Science from Hansei University, Korea, in

2010. Dr. Suh joined the faculty of the Department of Computer Science at Kyungin Women's University, Incheon, Korea, in 1998. He is currently a Professor in the Department of Tax Accounting, Kyungin Women's University. He is interested in network, operating system and computer graphics.