

# 스마트공장을 위한 빅데이터 애널리틱스 플랫폼 아키텍처 개발

신승준<sup>†</sup>, 우정엽<sup>\*\*</sup>, 서원철<sup>\*\*\*</sup>

## Developing a Big Data Analytics Platform Architecture for Smart Factory

Seung-Jun Shin<sup>†</sup>, Jungyub Woo<sup>\*\*</sup>, Wonchul Seo<sup>\*\*\*</sup>

### ABSTRACT

While global manufacturing is becoming more competitive due to variety of customer demand, increase in production cost and uncertainty in resource availability, the future ability of manufacturing industries depends upon the implementation of Smart Factory. With the convergence of new information and communication technology, Smart Factory enables manufacturers to respond quickly to customer demand and minimize resource usage while maximizing productivity performance. This paper presents the development of a big data analytics platform architecture for Smart Factory. As this platform represents a conceptual software structure needed to implement data-driven decision-making mechanism in shop floors, it enables the creation and use of diagnosis, prediction and optimization models through the use of data analytics and big data. The completion of implementing the platform will help manufacturers: 1) acquire an advanced technology towards manufacturing intelligence, 2) implement a cost-effective analytics environment through the use of standardized data interfaces and open-source solutions, 3) obtain a technical reference for time-efficiently implementing an analytics modeling environment, and 4) eventually improve productivity performance in manufacturing systems. This paper also presents a technical architecture for big data infrastructure, which we are implementing, and a case study to demonstrate energy-predictive analytics in a machine tool system.

**Key words:** Smart Factory, Big Data, Data Analytics, Machine Learning, Manufacturing Execution System, Energy Prediction

### 1. 서 론

최근 세계 제조강국인 독일, 미국, 일본, 중국 및 한국을 중심으로 제조 일류화 실현을 위한 경쟁이

치열하게 전개 중이다 [1]. 이러한 경쟁의 중심에는 제조기술과 첨단 정보통신기술간의 융합을 통해 제조공정의 생산성 향상을 달성하기 위한 제조혁신 필수 전략인 스마트공장이 있다 [2]. 스마트공장은 “제

※ Corresponding Author: Seung-Jun Shin, Address: (48547) 365 Sinseon-ro, Nam-gu, Busan, Korea, TEL: +82-51-629-5646, FAX: +82-51-629-5659, E-mail: sjshin@pknu.ac.kr

Receipt date: Jul. 28, 2016, Approval date: Aug. 9, 2016

<sup>†</sup> Graduate School of Management of Technology, Pukyong National University

<sup>\*\*</sup> Information Technology Laboratory, National Institute of Standards and Technology  
(E-mail: jungyub.woo@nist.gov)

<sup>\*\*\*</sup> Division of Systems Management and Engineering, Pukyong National University  
(E-mail: wcseo@pknu.ac.kr)

※ This research was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP) (No. 2016R1C1B1008820). This research was also supported by the Pukyong National University Research Fund (No. C-D-2015-1267).

품의 기획·설계, 생산, 유통·판매 등 전 과정을 정보통신기술로 통합, 최소 비용·시간으로 고객맞춤형 제품을 생산하는 미래형 공장”으로 정의될 수 있으며 [3], 인터넷으로 연결된 생산설비들이 스스로 모든 정보를 실시간으로 획득하여 자율적이고 협업적인 생산을 수행함으로써 최적화된 동적 생산의 구현을 가능하게 한다.

스마트공장을 위한 구현물은 현장 엔지니어가 활용하거나 지원받는 소프트웨어 형태의 도구로 제공될 것이다. 이 중에서 첨단화·지능화된 제조실행시스템(MES: Manufacturing Execution System)은 제조공정의 효율적 관리 및 개선을 위한 핵심도구가 될 것이다. 특히, 제조현장 데이터의 실시간 분석을 통한 제조공정의 진단, 예측과 최적화 달성을 가능하게 하는 데이터 애널리틱스(Data Analytics) 기능을 MES에 탑재함으로써, 경험에 의존하고 더디게 진행되던 의사결정 과정을 데이터에 근간한 실시간 의사결정 과정으로의 변환이 가능해진다. 여기서, 데이터 애널리틱스는 데이터로부터 유의미한 정보 및 지식을 추출하기 위한 비즈니스 인텔리전스의 한 분야로써, 추출되는 정보·지식의 대상 및 수준에 따라 묘사적 분석, 진단적 분석, 예측적 분석 및 지식적 분석으로 진화될 수 있다 [4]. 데이터 애널리틱스를 통해서 제조시스템에 대한 다양한 핵심성과지표(Key Performance Index)를 대상으로 멀티 스케일의 실시간적 진단, 예측 및 최적화 모델 개발을 가능하게 한다. 또한, 알려져 있지 않던 제조 오류나 문제점을 데이터를 통해 인과관계를 밝혀냄으로써, 문제 해결 및 성능 향상을 달성할 수 있다.

한편, 데이터의 수집, 처리와 활용은 데이터 애널리틱스의 핵심이다. 그러나, 아래와 같은 제조 데이터의 특성은 데이터 애널리틱스의 적용을 어렵게 한다. 첫째, 인간-기계-자재(Man-Machine-Material)으로 구성되는 제조객체(manufacturing objects)는 많은 양의 데이터를 생성하고 사용한다 (Volume 이슈). 둘째, 제조데이터는 시간, 데이터 소스 및 형식의 다양성이 존재한다 (Variety 이슈). 셋째, 제조현장에서의 실시간 의사결정을 위한 빠른 데이터 처리 및 응답이 이루어져야 한다 (Velocity 이슈). 기존의 관계형 데이터베이스 시스템으로는 이러한 제조데이터 특성에 의한 데이터 처리 요구성능의 만족이 용이하지 않기 때문에, 이를 해결하기 위한 분산형

데이터베이스 시스템과 제반 데이터 처리기술들, 소위 빅데이터 인프라스트럭처(이하, 빅데이터 인프라)와 융합된 데이터 애널리틱스, 즉 빅데이터 애널리틱스(Big Data Analytics)는 최적의 생산계획 및 제어를 위해 시기적절하고 정확한 의사결정을 가능하게 하는 스마트공장의 필수기술이다.

본 논문에서는, 스마트공장 실현을 위한 제조시스템용 빅데이터 애널리틱스 플랫폼 아키텍처를 제안한다. 제안되는 플랫폼 아키텍처는 소프트웨어 구현을 위한 개념구조로서, 제조설비에 특화된 데이터 애널리틱스 모델의 생성 및 적용에 필요한 기능들을 정의한다. 또한, 데이터 수집·저장·처리, 애널리틱스 모델링 및 제조 어플리케이션을 아우르는 데이터 애널리틱스와 빅데이터 인프라의 통합구조를 정의한다. 나아가, 상호호환적인 데이터 교환 환경을 고려하여 표준 제조데이터 인터페이스를 활용하며, 비용효과적인 구현 환경을 고려하여 오픈소스 솔루션을 활용한다. 제안되는 플랫폼 아키텍처의 개발물은 MES로의 내재화 혹은 MES와 통합화되어 데이터에 근간한 예측 및 최적화 의사결정을 수행하도록 지원한다.

본 논문의 구성은 다음과 같다. 2장에서는 관련된 국내외 연구동향을 서술한다. 3장에서는 빅데이터 애널리틱스 플랫폼 아키텍처를 제안한다. 4장에서는 구현중인 빅데이터 인프라를 설명하고, 데이터 애널리틱스를 적용한 공작기계 시스템의 전력에너지 예측 실험을 서술한다. 5장에서는 결론을 맺는다.

## 2. 국내외 연구동향

본 장에서는 제안되는 기술과 관련된 국내외 연구동향을 분석하고, 현재 기술수준 극복을 위한 플랫폼 아키텍처의 개념적 요구사항을 도출한다.

### 2.1 스마트공장

독일에서는 스마트제조(정보통신기술과 제조기술의 융합을 추구하는 제조 패러다임. 스마트공장은 스마트제조 구현결과물로 볼 수 있음)의 실현을 위하여 인더스트리 4.0(Industrie 4.0)이라 일컫는 민·관·학·연 거대 프로젝트를 정부주도로 추진 중이다 [6]. 인더스트리 4.0의 핵심기술은: 네트워크에 연

결된 자율적 협업을 실시하는 M2M(Machine-to-Machine), 네트워크를 통해 얻은 빅데이터의 활용, 생산부문의 MES와 ERP(Enterprise Resource Planning)-SCM(Supply Chain Management)-PLM(Product Lifecycle Management) 정보시스템간의 연결이다 [6]. 미국에서도 첨단제조 능력 강화를 통한 제조업의 국가경쟁력 부활을 목표로 민·관·학·연 프로젝트를 수행중이다 [1]. 첨단제조 중심에는 스마트공장이 있으며, 핵심기술로는: 빅데이터 애널리틱스, 사이버-물리 시스템, 무선 플랫폼, 스마트로봇, 3D프린팅 활용 적층제조, 상호운용성 및 표준 개발 등이 있다 [1]. 그 외에도 중국의 '제조강국 2025', 일본의 '산업재흥플랜' 등이 있으며 이러한 국가적 전략은 스마트공장의 실현으로 모아지고 있으며, 제조분야의 핵심 시스템인 MES의 고도화 및 첨단화를 실현 전술로 내세우고 있다 [6].

한편, 한국에서도 스마트공장추진단을 중심으로 스마트공장 보급사업을 적극 추진중이다 [6]. 추진단에서는 스마트공장 수준을 현장자동화, 공장운영, 기업자원관리 등의 기준에 따라 '정보통신기술 미적용', '기초수준', '중간수준1', '중간수준2' 및 '고도화'로 구분하며, 업종별 참조모델을 제공중이다 [6]. 스마트공장 요소기술은: 센서를 통해 데이터를 수집하여 플랫폼으로 전송하는 사물인터넷 기반 디바이스 기술, 디바이스에서 수집된 데이터를 처리, 분석 및 모델링하여 상위 어플리케이션에게 전달하는 플랫폼 기술, 플랫폼상에서 실행되는 MES, ERP, PLM 등의 제조 어플리케이션 기술로 구분된다 [7].

■ 요구사항: 스마트공장 실현을 위해서는 이기종 제조설비로부터 센서데이터를 수집, 저장 및 처리할 수 있는 플랫폼 기술이 필요하다. 또한, 이러한 플랫폼과 연계되어 기존에는 수행이 어려웠던 데이터 기반의 핵심성과지표의 예측 및 최적화 제어·계획을 실행하는 제조 어플리케이션 통합 기술의 개발이 필요하다.

## 2.2 제조실행시스템(MES)

MES는 공장의 실시간 모니터링, 작업의 계획 및 수행, 품질관리 등을 수행하는 정보시스템을 의미한다. 최근, 스마트공장의 핵심 정보시스템인 MES 고도화 및 첨단화를 위하여 연구 역량을 집중하고 있다. 국외의 경우, 빅데이터 애널리틱스 선도업체와

의 제휴 또는 연합 전략을 구사하며, 막대한 비용과 인력을 들여 킬러 어플리케이션화를 시도하고 있다. 국내의 경우, 대기업을 중심으로 대규모 고급인력을 활용하여 패밀리기업용 맞춤형 MES를 공급하고 있으며, 고성능 및 고가의 데이터 애널리틱스 기능을 제공중이다. 한편, 중소기업을 위한 보급형 MES 개발이 출현하고 있으나, 가격요건 제약으로 인하여 보급형 MES는 상급 언급된 기본기능 제공에 머무르고 있다. 그리고 국내의 MES 소프트웨어 수준은 최고 기술국 대비 70% 수준을 확보한 것으로 조사되고 있다 [8].

■ 요구사항: 기술수준의 격차 해소를 위하여 빅데이터 애널리틱스와 통합을 통한 지능형 MES 기술 개발이 필요하다. 특히, 중소기업을 대상으로 지능형 MES의 보급 노력이 필요하다. 이를 위해서는, 특정 벤더에 종속적이지 않은 데이터 인터페이스를 활용하고 저비용으로 시스템 구현과 통합을 달성하는 표준 및 오픈소스 기반의 빅데이터 애널리틱스 플랫폼 개발이 필요하다.

## 2.3 제조분야의 빅데이터 애널리틱스

빅데이터 애널리틱스는 IT산업분야의 킬러 어플리케이션으로 자리매김하고 있다. 일반적으로 공공분야, 의료분야, 금융, 통신, 상거래 산업과 같이 대규모 데이터가 생성되는 분야에서 주로 적용중이다 [9]. 제조업 분야에서도 꾸준한 빅데이터 애널리틱스와 MES간 통합을 위한 시도가 있었다. 그러나, 산업분야별 데이터의 생성량, 형태, 사용처, 프로토콜, 무결점 제어 등의 복잡성 및 다양성이 존재하는 제조업 특성상 빅데이터 애널리틱스의 활용은 쉽지 않다. 특히, 중소제조업체는 고비용과 전문인력 부족으로 인하여 통합작업의 어려움이 존재한다. 더 심각한 문제는 국내 중견·중소기업의 경우 66.7%가 '정보통신기술 미적용'이라고 응답한 바와 같이, 정보통신기술 자체가 적용 안 된 공장들이 많이 존재한다는 것이다. 특히, 뿌리산업의 경우는 제조공정의 정보통신기술화가 더욱 낮은 것으로 조사되고 있다 [8].

■ 요구사항: MES와 빅데이터 애널리틱스 통합은 제조업 분야와 IT산업 분야와의 융합을 의미하여, 제조 경쟁력 및 제조공정의 스마트화를 위해 필수적으로 이루어져야 한다. 이를 통해서 수율 및 품

질 향상, 제조공정의 가시성 및 유연성 확보, 비용 절감 그리고 장비가동을 향상과 같은 효과를 달성할 필요가 있다.

### 3. 빅데이터 애널리틱스 플랫폼 아키텍처

본 장에서 제안되는 빅데이터 애널리틱스 플랫폼은 단순히 제조분야로의 빅데이터와 데이터 애널리틱스의 적용을 위함이 아니다. 본 플랫폼의 목적은 빅데이터와 데이터 애널리틱스를 도구삼아 제조현장에서의 데이터 기반 의사결정을 통한 제조 지능화(manufacturing intelligence)를 구현하기 위함이다. 이를 위하여, 플랫폼 아키텍처를 제안하며, 세부사항을 설명한다.

Fig. 1은 제조시스템을 위한 빅데이터 애널리틱스 플랫폼 아키텍처를 표현한다. 이는 2장의 개념적 요구사항을 만족하도록 설계되었다(상세 요구사항은 하위 각 절에서 설명). 본 플랫폼을 통해 다음의 메커니즘을 달성하도록 한다: 1) 실제 공장을 에이전트 기반 공장으로 가상화, 2) 에이전트와 빅데이

터간 제조데이터 교환, 3) 제조데이터의 효율적 가공 및 처리, 4) 애널리틱스 모델 생성, 5) 애널리틱스 모델의 저장 및 관리, 6) 제조 어플리케이션과의 연계, 7) 물리적 또는 가상적 공장으로서의 실시간(정시) 의사결정 피드백. 이를 위하여, 본 플랫폼은 에이전트 기반 가상 공장, 빅데이터 인프라, 애널리틱스 모델링 어플리케이션, 제조 어플리케이션 레이어로 구성된다. 하위 절은 각 레이어의 상세 요구사항 및 기능을 기술한다.

#### 3.1 에이전트 기반 가상 공장

제조시스템에서 널리 쓰이는 전통적인 중앙집중 제어 구조는 수요량의 변동 및 변화에 대한 유연성이 부족한 단점을 가진다. 또한, 공장의 이상상황 발생에 대한 대처가 느리고, 제조시스템의 확장성과 재구성성 능력이 떨어진다 [10]. 이러한 단점들은 스마트공장에서 지향하는 자율적이고 협업적인 생산 환경 구축에 걸림돌로 작용하고 있다. 따라서, 제조시스템의 제어구조는 보다 자율적이고 재구성 가능

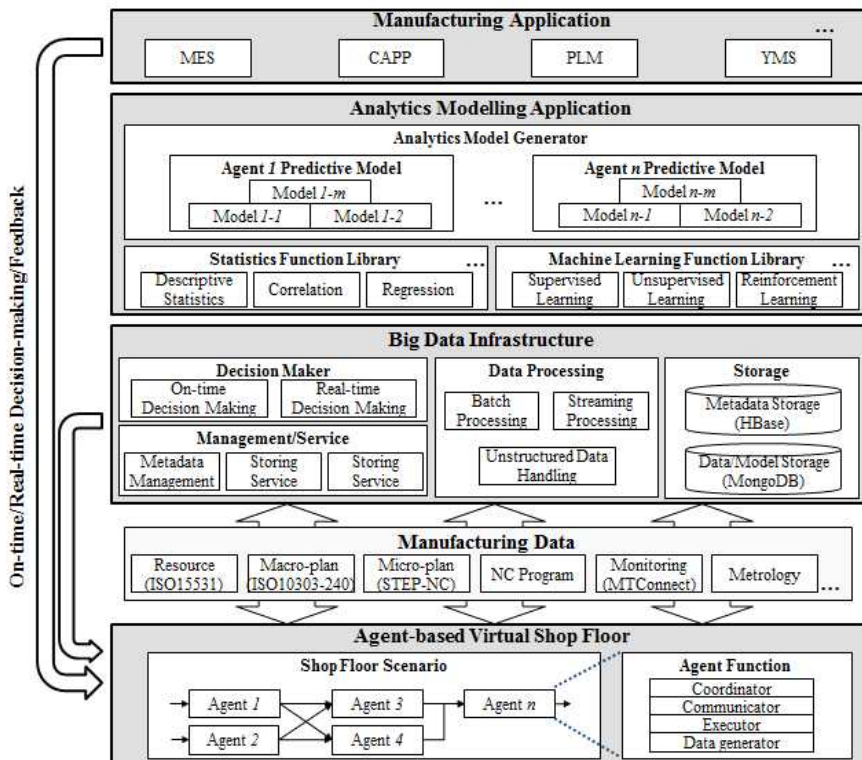


Fig. 1 A big data analytics platform architecture for manufacturing systems.

하고 선행적이고 상호작용적인 제어구조로의 변환이 필요하다. 이러한 제어구조는 에이전트 시스템을 통해 달성 가능하다. 에이전트 시스템이란 동적인 환경에서 작동하고, 자율적이고 지능적인 행위를 수행할 수 있는 컴퓨터 시스템을 의미하며 [11], 제조 시스템내에서는 제조객체를 가상객체 형태의 에이전트를 통하여 제조객체의 조정, 통신, 제어 및 정보 교환 역할을 수행하는 것이다.

여태까지는, 물리적인 제조객체와 가상 에이전트 간 정보 연결의 어려움으로 인해 제조시스템에서의 에이전트 시스템 구현이 어려웠다. 또한, 일반적으로 제조데이터는 폐쇄적이고, 벤더 의존적이고, 연결이 어려웠기 때문에 물리 제조객체로부터의 데이터 접근 및 수집이 어려웠다. 그러나, 최근 정보통신 기술의 발달로 인하여 물리 제조객체로부터의 효율적인 데이터 접근 및 활용이 가능해졌다. 특히, 정보통신기술분야의 사물인터넷, 사이버-물리 시스템 기술을 적용하여 생산설비에 임베디드된 데이터 수집장치와 인터넷 연결을 통해 물리 제조객체와 가상 에이전트 연결을 가능하게 해준다. 제조기술분야의 MTConnect [12]와 같은 표준형, 개방형 데이터 인터페이스는 제조설비로부터의 용이한 데이터 수집을 가능하게 해준다. 따라서, 제조시스템으로의 에이전트 시스템 적용 실현성은 보다 높아졌다.

Fig. 2는 에이전트기반 가상 공장의 유즈케이스

(Use Case)를 나타낸다. 이 가상 공장은 물리적 공장의 정보 허브로서, 에이전트들과 웹서비스간의 연결을 통해 다양한 어플리케이션과의 1-to-N 통합을 가능하게 한다. 또한, 물리적 제조객체를 식별하고 표현하며, 다른 객체와 통신하며, 물리적 제조객체로부터 데이터를 수집하고, 제조객체의 의사결정 및 제어를 수행한다. 빅데이터 애널리틱스와 연결되어 제조설비의 예측 모델에 기반하여 그들의 의사결정 및 제어를 수행하도록 예측 제어를 지원한다. 또한, 국소적으로는 제조설비 한 대에 대한 최적의 공정계획 및 제어파라미터를 결정하고, 광역적으로는 제조 시스템 전체에 대한 최적의 공정계획 및 자원할당 등을 결정하는 최적 제어를 지원한다. 만약, 예기치 않은 생산 스케줄 변화 발생시에는 제조설비간의 협의의 통해 대안의 최적 생산계획 및 할당을 실시하게 된다.

### 3.2 빅데이터 인프라스트럭처

1장에서 언급한대로, 제조데이터의 3V(volume, variety, velocity) 이슈가 존재한다. 구체적으로, 설비모니터링 데이터는 기계의 움직임과 활동을 시계열 형태로 표현한 것으로써, 통상적으로 수십분의 일 또는 수 초(sec) 단위로 측정되므로 대용량의 데이터가 생성된다. 공정계획(process planning) 프로그램이나 수치제어(NC: Numerical Control) 프로그램

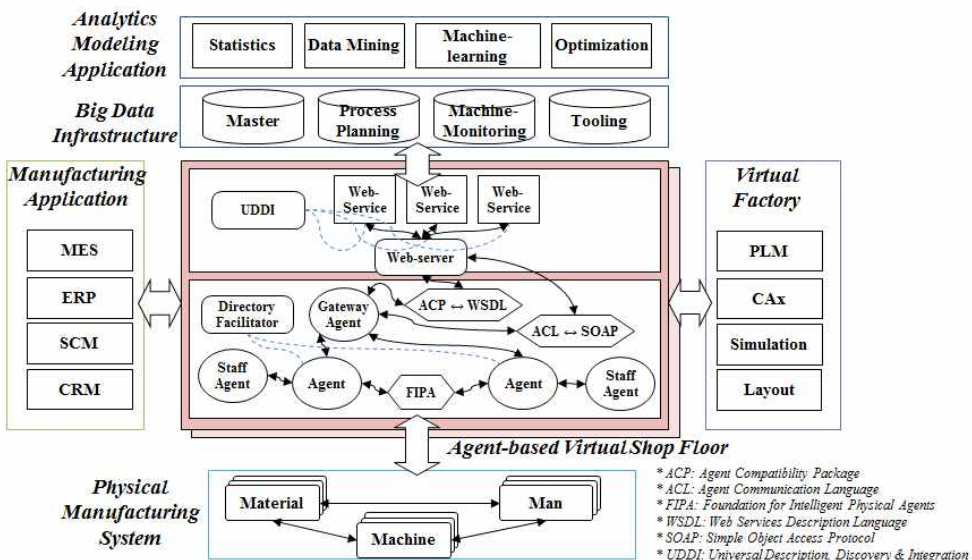


Fig. 2 Use case of agent-based virtual shop floor.

램(제조설비를 제어하는 언어)은 시간에 독립적인 형태이다. 반면, 설비모니터링 데이터는 시간 의존적이며 이들간의 구조 및 표현형태 또한 상이하다. 또한, 설비들은 실시간 제어 및 이상상황 감지를 실시해야 하므로 그들의 의사결정은 실시간으로 이루어진다. 따라서, 이러한 대용량의 다양한 제조데이터를 빠른 시간내에 핸들링할 수 있는 즉, '빅데이터 처리능력(big data capability)'을 갖춘 데이터 인프라 개발이 필요하다.

제조자(manufacturer)들은 제조데이터가 어디에 위치해 있든 어느 어플리케이션에서 생성되었든 상관없이 쉽게 제조데이터에 접근하고 이용하길 원한다. 또한, 제조자들은 제조데이터가 올바른지 그리고 신뢰하는 데이터소스로부터 수집되었는지에 대한 데이터 품질의 보장을 원한다. 따라서, 접근가능하고 신뢰할 수 있는 그리고 데이터 불확실성의 정량화가 되어 있도록 '데이터 투명성(data transparency)'을 갖춘 데이터 인프라 개발이 필요하다. 한편, 1장에서 언급한대로, 상호호환적인 데이터 교환 환경의 구현을 위해 '데이터 표준화(data standardization)'가 필요하다. 표준화의 결여는 제조설비 데이터의 획득 실패가 발생하거나 특정 기간·설비에서의 불완전하거나 부실된 데이터가 발생할 수 있기 때문이다.

Fig. 3은 빅데이터 인프라의 유즈케이스를 나타낸다. 이 인프라는 가상 공장으로부터 원시 제조데이터(raw manufacturing data)와 이와 관련한 구성데이터(제조데이터 획득 시점·대상·방법·형태·방식·원천 등의 마스터 정보)를 수집한다. 이 인프라는

세 개의 참여자(에이전트기반 가상 공장, 제조 어플리케이션, 애널리틱스 모델링 어플리케이션)와 데이터를 교환하는데, 획득된 데이터는 참여자에 따라 다양한 방식으로 활용된다. 가상공장의 경우, 획득된 데이터를 활용하여 공장의 상태를 모니터링하고 실시간·정시 제어를 실시한다. 여기서, 실시간 제어(real-time control)는 간단하고 반복적인 진단을 통하여 데이터 이상치가 발견된 설비에 적절한 제어를 실시하는 것을 의미한다. 예를 들면, 어떤 설비의 공정능력(process capability)에 이상치가 발견된 경우 해당 설비 가동을 멈추는 것이다. 정시제어(on-time control)는 시간허용치가 존재하여 다른 두 참여자에서 결정된 계획 및 제어 요소들을 다시 가상 공장으로 피드백하는 것을 의미한다. 제조 어플리케이션의 경우, 실시간각 변동하는 설비의 조건(condition)이나 상태(health) 데이터를 전달하여 제조 어플리케이션이 핵심성과지표의 모니터링 및 진단, 이상상황 경보 및 제조공정 제어 등을 실시하도록 한다. 이와 같이, 제조 어플리케이션은 그들의 고유목적인 단기(실시간이 아닌) 및 중기 의사결정에 집중할 수 있다. 한편, 애널리틱스 모델링 어플리케이션의 경우, 획득된 데이터는 제조설비에 커스터마이징되고 특화된 진단, 예측 및 최적화 모델 생성을 가능하게 한다. 제조데이터는 일반적으로 공정계획, 제어 파라미터 등은 원인인자, 설비모니터링 데이터는 결과인자 역할을 하므로 이 데이터들의 수집과 분석을 통하여 인과관계를 탐색하거나 발견할 수 있기 때문이다.

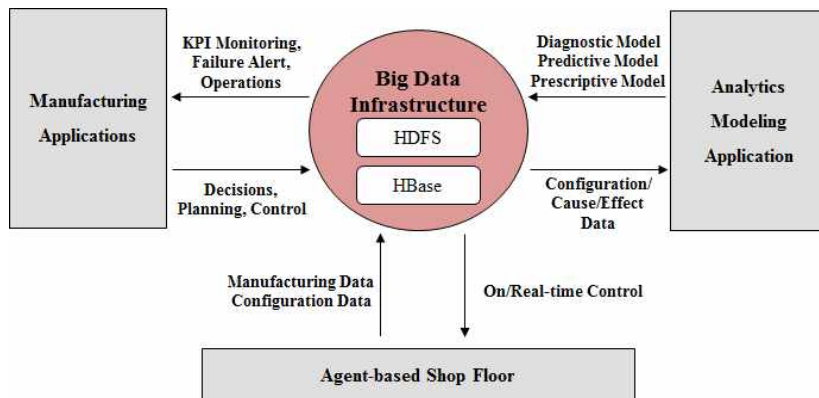


Fig. 3 Use case of big data infrastructure.



3.3 데이터 애널리틱스 어플리케이션

데이터 애널리틱스 어플리케이션의 목적은 획득된 제조데이터를 사용하여 애널리틱스 모델을 생성하기 위함이다. 즉, 이론적 또는 실험적 방식이 아닌, 과거에 진행되었던 제조공정으로부터 수집되고 축적된 데이터(historical data)를 기반으로 기계학습(machine-learning) 기법을 적용하여 진단, 예측 및 최적화 모델을 생성하는 것이다. 이러한 방식은 가정이나 국한적인 실험 데이터에 의존하지 않고, 주어진 공정구성(manufacturing configuration, 공정수행시 사용된 설비, 공구, 소재, 부가물질 등의 조합)에 따른 정확한 진단력, 예측력 및 최적화 능력을 제공한다. 예를 들면, 두 개의 서로 다른 공작기계 시스템은 각기 다른 성능을 내기 때문에 예측 모델을 표현하는 수학적 모형의 계수들이 상이할 수밖에 없다. 이 때, 제안된 방식을 통하여 각 공작기계에서 수집된 데이터를 바탕으로 각기 다른 예측 모델의 모형 계수들을 결정함으로써 정확한 예측력을 확보할 수 있다.

Fig. 4는 데이터 애널리틱스 어플리케이션에 실행되는 애널리틱스 모델 생성을 위한 논리적인 절차를 나타낸다. 아래의 단계는 각 절차를 서술하며, 상세 설명을 위하여 절삭공정용 공작기계 시스템의 전력에너지 예측을 대상으로 서술한다(4.2절은 이에 대한 실험을 서술). 여기서, 컴포넌트 모델은 단수

개의 특정 공정구성에 대해서 원인과 결과 데이터간 상관관계를 수치적으로 규명하여 목표 핵심성과지표를 예측할 수 있는 모델을 의미한다. 만약 공정구성 속성값 모두 같을 경우, 같은 컴포넌트 모델이 사용된다. 반면, 공정구성 속성값이 다르면, 다른 공정구성에서 가공된 것이므로 다른 컴포넌트 모델을 사용해야 한다.

1. 공정속성 데이터 정의(Process Attribute Data Identification): 수집되어야 할 속성 데이터를 정의하는 단계이다. 공정속성은 상기 언급한 공정구성을 정의하기 위해 수집되는 속성뿐만 아니라, 원인-결과 데이터셋 구성을 위한 원인속성과 결과속성으로 구성된다. 먼저, 하나의 컴포넌트 모델이 적용될 하나의 특정 공정상황을 특성화하기 위한 공정구성 데이터를 수집해야 한다. 절삭공정에서의 공정상황은 공작기계, 가공소재, 공구, 절삭유, 공정종류, NC코드 및 코드유형으로 정의할 수 있다. 그리고, 원인 및 결과 데이터를 구성하는 속성값의 정의가 필요하다. 절삭공정에서는 원인 속성을 공작기계 제어를 위한 파라미터-이송속도(feedrate), 스피indle 회전속도(spindle speed), 절삭깊이(cutting depth), 절삭폭(cutting width)-로 정의할 수 있다. 한편, 결과속성은 공작기계의 전력에너지(electrical energy)로 정의한다. 이는 제어 파라미터에 따라 전력에너지는 가변하므로, 둘 간의 기계학습 모형을 생성하여 에너지 예측이 가능하기 때문이다.

2. 데이터 전처리(Data Pre-processing): 컴포넌트 모델의 신뢰성 확보 및 불확실성 제거를 통하여 고품질의 데이터를 획득하기 위한 데이터 처리 단계이다. 기계학습은 컴퓨터가 직접 훈련데이터로부터 지식을 학습하여 문제를 푸는 기법이므로, 훈련데이터를 필요로 한다. 나아가, 기계학습에 의한 모형은 훈련데이터의 품질에 많은 영향을 받으므로, 데이터 클리닝(cleaning), 정규화(normalization), 변형화(transformation) 및 이상치(outlier) 제거 등을 포함하는 데이터 전처리 과정을 통한 고품질의 훈련데이터 생성이 필요하다 [13]. 이를 통하여, 중요 정보의 손실이 없으면서 동시에 비신뢰적이거나 노이즈가 존재하는 데이터를 제거하여 데이터 불확실성을 줄일 수 있다.

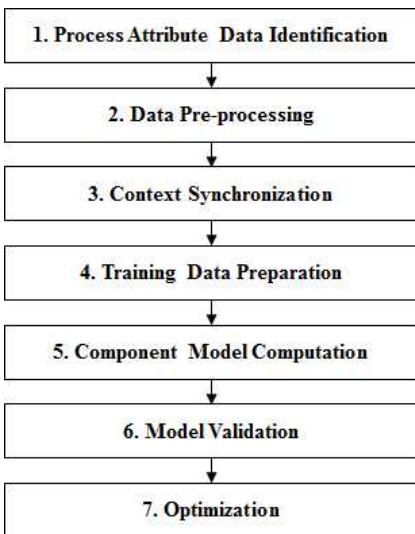


Fig. 4 Procedure of analytics modeling.

3. 상황 동기화(Context Synchronization): 훈련 데이터 생성을 위하여 각기 다른 소스에서 수집된 공정구성, 원인 및 결과 데이터를 시간에 따라 동기화(synchronization)하는 단계이다. 설비모니터링 데이터는 시간에 따른 설비의 움직임과 행동을 관찰하고 측정된 스냅샷(snapshot)의 연속으로 볼 수 있다. 이러한 설비모니터링 데이터의 분석은 해당 설비가 왜 그러한 움직임과 행동을 취했는지에 대한 상황을 이해하는데 활용될 수 있다. 이 때, 설비모니터링 데이터는 결과 데이터이므로, 이와 함께 원인 데이터인 공정계획 및 NC 데이터와 함께 분석이 필요하다. 왜냐하면, 설비의 움직임과 행동은 상기 원인 데이터에 의해 결정되기 때문이다. 따라서, 설비모니터링 데이터는 해당 공정계획 및 NC 데이터와 정합(matching)시키는 것이 필요하다. 즉, 상황 동기화는 시간에 따른 결과 데이터와 대응되는 원인 데이터를 정합시키는 과정이다.

4. 훈련데이터 준비(Training Data Preparation): 상황 동기화된 데이터셋을 얻었다면, 공정구성별 훈련 데이터셋을 생성하는 단계이다. 이는 다음 단계에서 훈련 데이터로부터 컴포넌트 모델을 산출해야 하기 때문이다. 하나의 훈련 데이터셋은 공정구성, 원인 및 결과 데이터의 집합으로 구성된다. 이 때, 공정구성 데이터는 한 개의 전체 데이터셋에서 복수 개의 공정상황별 훈련 데이터셋들로 분리하기 위한 분류기준으로 사용되며, 한 개의 훈련 데이터셋은 하나의 컴포넌트 모델을 만들기 위해 사용된다.

5. 컴포넌트 모델 산출(Component Model Computation): 훈련 데이터셋의 학습을 통하여 수학적 모형으로 표현되는 컴포넌트 모델을 생성하는 단계이다. 이 단계는 기계학습 또는 회귀 기법에 의해 훈련 데이터의 학습이 가능하다. 예를 들어, 인공신경망(Artificial Neural Network) 기법을 적용할 수 있다. 인공신경망으로부터 원인데이터(제어파라미터: 이송속도, 스핀들 회전속도, 절삭깊이, 절삭폭)와 결과데이터(에너지)간의 관련성을 수학적 모형으로 산출할 수 있다. 하나의 공정구성은 이러한 수학적 모형을 최소 한 개 이상 갖게 되어, 어떤 원인속성 값들을 입력했을 때, 예측 결과속성 값이 이 모형에 의해 계산되는 것이다. 식 (1)은 이를 수학적 모형으

로 표현한 것이다. 여기서,  $x$ 는 원인속성 변수,  $y$ 는 결과속성 변수,  $p$ 와  $q$ 는 각 레이어의 뉴런 개수,  $w_{oj}$ 와  $w_{ji}$ 는 각 레이어에 존재하는 각 뉴런의 가중치값,  $f_o$ 와  $f_h$ 는 시그모이드 함수를 의미한다.

$$y = f_o \left( \sum_{j=0}^p w_{oj} f_h \left( \sum_{i=0}^q w_{ji} x_i \right) \right) + \epsilon \quad (1)$$

6. 모델 검증(Model Validation): 산출된 컴포넌트 모델의 신뢰성 및 성능을 측정하는 단계이다. 즉, 컴포넌트 모델이 주어진 임계치를 만족하는 성능을 내는지를 정형화되고 정량적인 방법을 통해 측정하는 것이다. 다양한 모델 검증 방법이 존재하며 이를 적절히 활용하는 것이 필요하다. 예를 들어, 전체 데이터셋을 훈련 데이터셋과 테스트 데이터셋으로 분할하여 검증하는 데이터 분할 또는 교차확인 방법 등을 사용할 수 있다. 그리고, 예측값과 실제값간의 차이를 나타내는 평균제곱근오차(RMSE: Root Mean Square Error)를 계산하여 모델의 성능을 측정할 수 있다.

7. 최적화(Optimization): 컴포넌트 모델을 이용하여 주어진 공정상황에 대한 최적의 원인데이터 값(여기서는 제어파라미터) 및 해당 결과데이터 값을 계산하는 단계이다. 컴포넌트 모델 산출의 의미는 원인속성과 결과속성간의 관계를 수학적으로 표현할 수 있다는 것이며, 최적화 과정에 필요한 의사결정변수( $x$ )와 목적함수( $y$ ) 정의가 가능하다는 의미이기도 하다. 따라서, 제약조건(constraints) 정의와 함께 다양한 최적해 기법을 사용하여 최적의 의사결정변수와 목적함수 값을 산출할 수 있다. 예를 들어, 에너지를 최소로 사용하는 제어파라미터를 산출하되 표면조도의 임계치가 제약조건으로 주어진다면, 이를 최적해 탐색 과정을 통하여 최적 제어파라미터 및 최소 에너지 값과 표면조도 값의 산출이 가능하다.

### 3.4 제조 어플리케이션

MES, PLM과 같은 제조 어플리케이션은 어플리케이션 인터페이스를 통해 플랫폼과 커뮤니케이션을 실행한다. 사용자는 이러한 제조 어플리케이션을 통해 플랫폼과 연계하며, 제조데이터 및 애널리틱스 모델을 사용하여 목적에 부합하는 진단, 예측 및 최



적화 등을 실시한다. 더불어, 제조상의 문제점들을 정형화하고 해당 애널리틱스 모델을 생성하여 이를 해결할 수 있다. 애널리틱스 모델을 기반으로 도출된 의사결정요소(공정계획 및 제어파라미터 등)는 실시간으로 또는 정시적으로 공장의 제조공정으로 피드백되어 적용된다. 예를 들어, 컴포넌트 모델에 의해 에너지 값이 예측된 제어파라미터를 동작기계 시스템으로 전송 및 적용하여 실제가공을 수행하게 된다.

4. 구현 및 실험

본 장에서는 절삭가공 시스템을 대상으로 구현 및 실험 사례를 서술한다. 4.1절은 구현중인 빅데이터 인프라의 기술구조도를, 4.2절은 제안된 애널리틱스 모델링 방식의 실현가능성(feasibility) 및 실용성(practicability) 확인을 위한 절삭가공에서의 에너지 예측을, 4.3절에서는 고찰을 서술한다. 본 논문에서는 3.3절에 언급한 최적화 단계는 배제하기로 한다.

4.1 빅데이터 인프라 구현

Fig. 5는 빅데이터 인프라의 기술구조도를 나타낸다. 3.2절에 언급한대로, 원시데이터 계층(Raw Data Layer)에서는 세 개의 참여자는 그들의 목적에 맞게 원시 제조데이터를 생성한다. 표준데이터

계층(Standard Data Layer)에서는 원시 데이터가 상호호환성 확보를 위하여 표준 데이터 포맷으로 변환된다. 제조분야에서는 다양한 데이터 포맷이 존재하므로, 이들을 언제 어디서나 활용가능하도록 단순하고 통일된 표준 데이터 포맷인 JSON(Java Script Object Notation) [14] 형식으로 변환한다. 예를 들어, 절삭가공 공정의 공정계획 데이터를 표현하는 국제표준인 ISO14649 형태의 데이터를 JSON 형식으로 변환함으로써, 빅데이터 인프라 내에서 해석가능하고 사용가능하도록 한다. 한편, 애널리틱스 모델의 표현은 데이터마이닝그룹에서 제정한 예측모델 표현방식의 표준인 PMML(Predictive Model Markup Language) [15]를 사용한다. Table 1은 빅데이터 인프라에 포함되는 각 기능과 기능구현을 위해 사용중인 오픈소스를 서술한다.

4.2 절삭가공의 에너지 예측

3.3절에 제시된 애널리틱스 모델링 방식에 의거하여, 한 파트의 가공을 위해 코딩된 하나의 NC 프로그램 실행에 따라 소비되는 전력에너지 예측을 위하여 실험을 실시하였다. Fig. 6은 실험을 위해 2.5차원 밀링가공(평면절삭을 위해 X, Y축 동시제어 가공을 실시하고, 한 면이 가공되면 Z축을 이동해 다시 동일평면 절삭을 반복하는 가공)을 통해 제조된 파트를, Table 2는 실험 환경을 나타낸다. Table 3과 같이 12번의 테스트를 실시한다. 12번의 테스트를

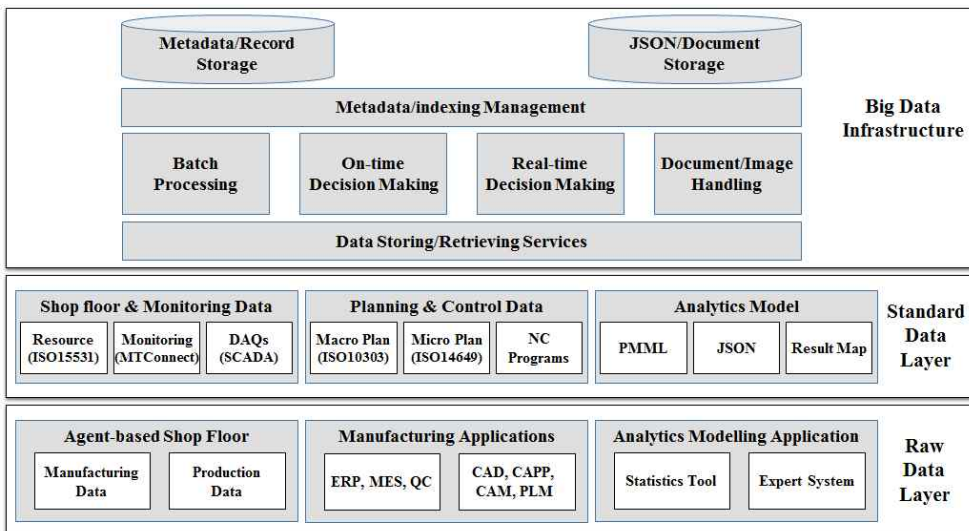


Fig. 5. A technical architecture of big data infrastructure.

Table 1. Functions and solutions for big data infrastructure

Function	Description	Solution
Data storing/retrieving service	Provide input/output interfaces to the participants	REST/AXIS Web-Server
Batch processing	Process large datasets with a parallel and distributed algorithm	MapReduce
On-time decision making	Support on-time control by manufacturing application and analytics modeling application	TBD (To be determined)
Real-time decision making	Support real-time control by big data infrastructure	TBD
Document/Image handling	Handle unstructured data formats including image and PDF files	TBD
Metadata/Indexing management	Generate and manage metadata to identify manufacturing configuration for data querying and retrieving	Hadoop Distributed File System
Metadata/Record storage	Store metadata associated with retrieving raw data from JSON/Document storage	HBase
JSON/Document storage	Store JSON-formatted raw manufacturing data	MongoDB



Fig. 6. A test part.

통하여, 각기 다른 제어파라미터-이송속도(feed-rate), 스펀들 회전속도(spindle speed), 절삭깊이(cutting depth)-와 측정된 전력에너지 값을 이용하여 에너지 예측 모델을 생성한다. 그 후, 예측 모델에 의한 예측 전력에너지 값과 실제 측정된 전력에너지 값과의 비교를 실시한다. 일반적으로, 제어파라미터는 소재·형상·공구에 따른 추천절삭 조건표에 주어진 범위내에서 가공 안전성을 고려하여 사용자가 적절하게 선정한다. 본 실험에서의 제어파라미터 선정은 실험계획법(Design of Experiments)에 의거하지 않고, 작업자가 추천절삭 조건 범위내에

Table 2. Setup of experiments

Property	Specification
Machine	Mori Seiki NVD 1500 DCG
CNC (Computerized Numerical Controller)	Fanuc 0i series
Workpiece	Material: Cold finish mild steel 1018 Volume: 10.16cm×10.16cm×1.27cm
Coolant	Emulsion type
Cutting tool	Solid carbide
Tool geometry	Type: Flat end mill Diameter: 8mm Number of flutes: 4
Power meter	System Insights High Speed

Table 3. Setup of process parameters

Test	Feedrate (mm/tooth)	Spindle speed (RPM)	Cutting depth (mm)
1	0.0127	1500	1.5
2	0.0127	2000	1.5
3	0.0127	1750	1
4	0.0229	1750	1
5	0.0127	1750	2
6	0.0178	1500	1
7	0.0178	2000	1
8	0.0178	2000	2
9	0.0178	1750	1.5
10	0.0076	1750	1.5
11	0.0152	1750	1.5
12	0.0127	1750	1.5

임의 값을 입력한 것으로 가정한다.

실제가공을 위해 코딩한 공정계획 데이터(ISO 14649 프로그램)와 NC 프로그래밍 데이터(NC 프로그램), 그리고 공작기계로부터 획득한 실제 설비 모니터링 데이터(MTConnect 문서)를 수집하였다. Fig. 6은 수집된 세 가지 데이터셋의 일부를 발췌한 예시이다. Fig. 6(a)는 ISO14649 프로그램, Fig. 6(b)는 NC 프로그램, Fig 6(c)는 MTConnect 문서로써, 각기 다른 표준에 의하여 각기 다른 형태를 가진다 (Variety 이슈). 또한, MTConnect 문서의 경우, 평균 0.365초 단위로 수집되었으며, 본 실험에서는 약 1.06GB 용량의 세 가지 데이터셋이 수집되었다 (Volume 이슈). 이러한 데이터는 JSON 형식으로 변환되어 Fig. 5의 JSON/Document storage에 저장된다. Fig. 7은 ISO14649 프로그램의 일부를 JSON 클래스 형식으로 변환한 것을 나타낸다. 3.3절에 언급한대로, 본 실험을 위해서 공정속성 데이터를 아래와 같이 정의하였으며, 상기 언급된 데이터를 이용하여 공정 속성 데이터 추출, 데이터 전처리 및 상황 동기화를 실시하였다.

공정구성 속성 = {공작기계(Machine), 가공소재(Workpiece), 절삭유(Coolant), 절삭공구(Cutting tool), 공정종류(Operation: contouring, slotting, pocketing or drilling), NC코드(Command: G01 or G02/G03), 코드유형(Trajectory: approach, re-

```
#1=PROJECT('EXECUTE EXAMPLE1',#2,(#4),$,,$,$,#10);
#2=WORKPLAN('MAIN WORKPLAN',(#30,#36,...),$,,$,$);
#4=WORKPIECE('CUBICAL WORKPIECE',#5,0.001,$,$,#80,());
#5=MATERIAL('CR-1018','STEEL',(#6));
#10=MACHINE_TOOL_SPECIFICATION('MILLING_MACHINE',#11,...);
#11=DEVICE_ID('MILLING_MACHINE1', 'NVD1500DCG', $, '$', $);
...
#36=MACHINING_WORKINGSTEP('OPEN SLOT1',#60,#3400,#500,$);
...
#500=BOTTOM_AND_SIDE_ROUGH_MILLING($,$,'OPEN_SLOT1',
10.0,$,#510,#530,#540,$,#550,#560,#570,2.0,8.0,0.0,0.0);
#510=MILLING_CUTTING_TOOL('ENDMILL',#511,(#513),80.0,$,$);
#530=MILLING_TECHNOLOGY(0.007,TCP,$,209.4,$,F.,F.,F.,$);
#540=MILLING_MACHINE_FUNCTIONS(T,$,$,F.,$(.),T,$,$());
#570=CONTOUR_PARALLEL(0.05,T,...,CW,...CONVENTIONAL.);
...
#3400=SLOT('SLOT1',#4,(#500),#3401,#3410,(),$,#3415,$,#3418,#3420);
```

(a) ISO14649 program

```
N95 G00
N96 X-.103 Y48.8
N97 Z2.
N98 G01 F25
N99 Z-1.5
N100 G41 X-6.071 D01 F74
N101 G01 Y31.8
N102 X-6.073 Y31.7
N103 G03 X1.848 Y31.7 I3.96 J-.099
N104 G01
N105 X1.846 Y31.8
N106 Y48.8
```

(b) NC program

```
...
<ComponentStream name="X_AXIS" component="Linear" componentId="x_axis">
<Position dataItemId="X_POSITION" ... timestamp="2014-12-11T22:57:58.835Z">-6.071</Position>
<Position dataItemId="X_POSITION" ... timestamp="2014-12-11T22:57:59.143Z">-6.071</Position>
...
<ComponentStream name="Y_AXIS" component="Linear" componentId="y_axis">
<Position dataItemId="Y_POSITION" ... timestamp="2014-12-11T22:57:58.835Z">-47.808</Position>
<Position dataItemId="Y_POSITION" ... timestamp="2014-12-11T22:57:59.143Z">-47.125</Position>
...
<ComponentStream name="Z_AXIS" component="Linear" componentId="z_axis">
<Position dataItemId="Z_POSITION" ... timestamp="2014-12-11T22:57:58.835Z">-1.500</Position>
<Position dataItemId="Z_POSITION" ... timestamp="2014-12-11T22:57:59.143Z">-1.500</Position>
...
<ComponentStream name="Systems" component="Electric" componentId="Systems">
<Wattage dataItemId="WATTAGE" ... timestamp="2014-12-11T22:57:58.835Z">2811.500</Wattage>
<Wattage dataItemId="WATTAGE" ... timestamp="2014-12-11T22:57:59.143Z">2884.900</Wattage>
```

(c) MTConnect document

Fig. 7. Examples of ISO14649, NC and MTConnect data-sets.

tract, linear feed or circular feed))

원인 속성 = {이송속도(feedrate), 스핀들 회전속도(spindle speed), 절삭깊이(cutting depth)}

결과 속성 = {electrical energy}

본 실험에서는 공정계획 및 NC 프로그래밍 데이터 분석으로부터 51개의 각기 다른 공정구성 조합이 존재함을 확인하였다. 이를 토대로, 51개의 에너지

```

"headerSection" : {
  "fileDescription" : {
    "description" : [ "EXAMPLE OF NC PROGRAMME FOR MILLING: COMPLEX DESIGN." ],
    "implementation_level" : "6:2"
  },
  "fileName" : {
    "name" : "MILLING_STEEL_1.nc",
    "time_stamp" : "2015-04-23T01:53:19.000Z",
    "authrs" : [ "SHIN", "MOO" ],
    "organization" : [ "NIST", "NIST" ],
    "preprocessor_version" : "STEP-NC Controller Ver1.0",
    "authorization" : null
  },
  "filePopulation" : null,
  "fileSchema" : {
    "schema_identifiers" : [ "MACHINING_SCHEMA", "MILLING_SCHEMA" ]
  },
  "sectionContext" : null,
  "sectionLanguage" : null
}
    
```

Fig. 8. An example of a JSON object class for ISO14649 program.

컴포넌트 모델을 생성하였다. 컴포넌트 모델 생성시에는 데이터 마이닝 툴인 KNIME에서 제공하는 Multilayer Perceptron Neural Network 함수 [16]를 이용하였다. Table 4는 12번의 테스트별 실제 측정된 전력에너지와 예측된 전력에너지 값을 나타내며, 두 값간 차이 비율 및 RMSE를 포함한다. 비교 결과, 가장 큰 차이 비율은 최대 1.08%로 나타났으며, 모두 1.08%이내의 예측 정확성을 보여주고 있다.

4.3 고찰

본 실험에서, 전력에너지의 높은 예측력이 나타나는 이유는 제안된 방식에 의거하여 공정구성을 코드유형 수준까지 세분화하여 컴포넌트 모델을 생성했기 때문이다. 코드유형 중 approach(접근) 및 retract(후퇴)는 절삭공구와 가공소재간 접촉이 존재하지 않는 에어커팅(air cutting)을 수행하는데, 이때 절삭력이 적용되지 않아 순수하게 공작기계의 작동에 소요되는 전력에너지(기계 turn-on, 공구이송,

스핀들 회전 등에 소요되는 에너지)만을 소비하게 된다. 반면, feed(보간)계열 코드유형은 절삭공구와 가공소재간 접촉이 발생하여 실제커팅(real cutting)이 적용되며, 이 때에는 공작기계 작동을 위한 전력 에너지뿐만 아니라 절삭력에 의한 전력에너지가 동시에 소요된다. 이처럼, 다른 성질의 전력에너지를 고려하여 코드유형 수준까지 세분화된 각기 다른 컴포넌트 모델을 생성함으로써, 측정값에 근접한 에너지 예측력을 갖출 수 있다.

그리고, 기존의 이론적 또는 실험적 방식에 대비하여, 제안된 방식은 실용성을 향상시킬 수 있다. 기존의 연구들은 단순히 제어파라미터와 에너지간의 상관성을 규명하는데 집중하였으므로, NC 프로그램 실행에 소비되는 전력에너지의 예측이 어려웠다. 특히, 실험적 방식은 실험계획법에 의거하여 단수 또는 소수 개의 공구이송경로만을 대상으로 제어파라미터에 의한 에너지 상관성을 예측력을 규명하는데 집중하여 많은 수의 공구이송경로가 존재하는 NC 프로그램 전체에 대한 에너지 예측이 어려웠다. 반면에, 제안된 방식은 실험계획법 없이 NC 프로그램의 실행에 따른 에너지 예측을 가능하게 할 수 있다.

5. 결 론

본 논문에서는, 스마트공장 실현을 위한 제조시스템용 빅데이터 애널리틱스 플랫폼 아키텍처를 제안하였다. 플랫폼 아키텍처는 제조설비에 특화된 데이터 애널리틱스 모델의 생성 및 적용을 가능하게 하기 위하여 데이터 수집·저장·처리, 애널리틱스 모

Table 4. Experiment results

Test	Measured Energy (kJ)	Predicted Energy (kJ)	Difference(%)	RMSE(J)
1	13,952.5	13,901.1	-0.37	28.67
2	11,382.1	11,414.2	0.28	29.72
3	19,535.2	19,457.8	-0.40	23.62
4	9,830.3	9,823.5	-0.07	28.39
5	9,943.1	10,007.9	0.65	34.50
6	13,365.9	13,417.1	0.38	25.30
7	11,044.0	11,076.5	0.30	26.26
8	6,012.6	5,947.6	-1.08	41.10
9	9,750.7	9,720.8	-0.31	32.58
10	19,281.6	19,298.4	0.09	21.82
11	10,791.6	10,873.6	0.76	30.12
12	12,580.1	12,529.8	-0.40	29.98

델링 및 제조 어플리케이션을 아우르는 데이터 애널리틱스와 빅데이터 인프라의 통합구조를 정의한 것이다. 또한, 구현중인 빅데이터 인프라의 기술구조를 소개하고, 실험을 통한 제안된 애널리틱스 모델링 방식의 실현가능성 및 실용성을 확인하였다.

제안된 플랫폼의 성공적인 실현은 제조분야에서의: 1) 빅데이터 애널리틱스를 통한 제조 지능화 획득, 2) 표준 데이터 인터페이스 및 오픈소스 솔루션 활용을 통한 비용 효과성 획득, 3) 한국형 스마트공장 '고도화' 달성을 위한 하나의 참조모델 역할, 4) 궁극적으로는 제조시스템의 생산성 향상을 달성할 것으로 기대한다.

그러나, 빅데이터 인프라는 구현 중이어서, 이로 인하여 데이터 애널리틱스 어플리케이션과의 통합 및 성능 검증이 완료되지 않았다. 또한, 데이터 불확실성에 대한 정량적 측정, 실제 운영중인 제조 어플리케이션으로의 통합 등이 수행되지 않은 점은 본 논문의 한계이다. 향후, 지속적인 빅데이터 인프라 구현을 통한 성능 검증, 데이터 불확실성의 정량화 방법론 개발, 데이터 애널리틱스 어플리케이션 및 제조 어플리케이션과의 통합을 위한 인터페이스 개발 등을 수행할 것이다.

## REFERENCE

- [ 1 ] KIAT, *Research Trend of Smart Manufacturing in U.S.A*, KIAT Industrial Technology Policy Brief 2014-21, 2015.
- [ 2 ] KOSF, *Industrial Reference Models for Smart Factory Propagation*, KOSF Report, v.2.0, 2015.
- [ 3 ] KSA, *Global Trends and Korean Standardization Strategies for Smart Factory*, KSA Policy Study Issue Paper 012, 2015.
- [ 4 ] Deloitte, 2016 Global Manufacturing Competitiveness, *National Competitiveness Forum*, 2015.
- [ 5 ] A. Pavlo, E. Paulson, A. Rasin, D.J. Abadi, D.J. DeWitt, S. Madden, et al, "A Comparison of Approaches to Large-scale Data Analysis," *Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data*, pp. 165-178, 2009.
- [ 6 ] G.T. Lee, G.J. Lee, and B.H. Song, *Technology Trends of Smart Factory*, KEIT PD Issue Report 15-4, 2015.
- [ 7 ] S.H. Baek, *The-state-of-the art and Standardization Strategies for Smart Manufacturing*, KISTEP Issue Paper 2016-03, 2016.
- [ 8 ] KDB, *The Feasibility of Smart Factory for Advancing Korean Manufacturing Industry*, KDB Industrial Issue Paper, 2015.
- [ 9 ] Y.M. Song, and C.S. Lee, "A Study on the Big Data Analysis System for Searching of the Flooded Road Areas," *Journal of Korea Multimedia Society*, Vol. 18, No. 8, pp. 925-934, 2015.
- [ 10 ] W. Shen, Q. Hao, H.J. Yoon, and D.H. Norrie, "Applications of Agent-based Systems in Intelligent Manufacturing: an Updated Review," *Advanced Engineering Informatics*, Vol. 20, No. 4, pp. 415-431, 2006.
- [ 11 ] L. Monostori, J. Vancza, and S.R.T. Kumara, "Agent-based Systems for Manufacturing," *CIRP Annals - Manufacturing Technology*, Vol. 55, No. 2, pp. 697-720, 2006.
- [ 12 ] MTConnect Institute, *MTConnect@ Standard Part 1 - Overview and Protocol*, The Association for Manufacturing Technology, version 1.3.0, 2014.
- [ 13 ] S. Kotsiantis, D. Kanellopoulos, and P. Pintelas, "Data Preprocessing for Supervised Learning," *International Journal of Computer Science*, Vol. 1, No. 2, pp. 111-117, 2006.
- [ 14 ] ECMA International, *The JSON Data Interchange Format*, ECMA-404, 1<sup>st</sup> edition, Geneva, Switzerland, 2013.
- [ 15 ] PMML 4.2 - General Structure, <http://dmg.org/pmml/v4-2-1/GeneralStructure.html>, Data Mining Group, (accessed Mar., 15, 2016).
- [ 16 ] KNIME Analytics Platform, <http://www.knime.org/knime-analytics-platform>, (accessed Mar., 14, 2016).



신 승 준

2002년 고려대학교 기계공학과  
공학사  
2005년 POSTECH 대학원 산업  
공학과 공학석사  
2010년 POSTECH 대학원 산업  
경영공학과 공학박사

2010년~2012년 삼성전자 생산기술연구소 책임연구원  
2013년 삼성SDS MES개발팀 책임컨설턴트  
2013년~2015년 미국 표준기술연구소(NIST) 객원연구원  
2015년~현재 부경대학교 기술경영전문대학원 조교수  
관심분야: 스마트 제조, 빅데이터 애널리틱스, 사이버-  
물리 제조시스템, 친환경 제조



서 원 철

2003년 POSTECH 산업경영공학  
과 공학사  
2010년 POSTECH 산업경영공학  
과 공학박사  
2010년~2012년 삼성전자 LCD사  
업무 책임연구원

2012년~2013년 한국지식재산연구원 부연구위원  
2013년~현재 부경대학교 시스템경영공학부 조교수  
관심분야: 특허마이닝, 기술지능, 기술 및 제품기회발굴,  
기술테마분석



우 정 엽

2001년 Postech 산업경영공학과  
공학사  
2003년 Postech 산업경영공학과  
공학석사  
2007년 Postech 산업경영공학과  
공학박사

2005년~2007년 미국 표준기술연구소 (NIST) 객원 연구원  
2007년~2008년 삼성 SDS 글로벌 ERP 책임컨설턴트  
2008년~2009년 ORNL 연구교수  
2010년~현재 NIST HL7 Testbed 분석 및 개발팀, 팀장  
관심분야: 스마트 제조, 빅데이터, Health IT, Software  
Testing, 사이버-물리 제조시스템