

제조 공정 분석을 위한 빅데이터 클라우드 서비스

Big data Cloud Service for Manufacturing Process Analysis

이용혁¹ · 송민석^{1*} · 하승진² · 백태현² · 손숙영²

울산과학기술원(UNIST) 경영공학과¹
현대중공업 안전·경영지원본부 공정지원부²

요약

정보통신 기술의 발달로 과거에는 다룰 수 없었던 대용량의 데이터 처리가 가능해지면서 빅데이터의 관심이 고조되고 있다. 제조 산업은 축적된 데이터가 풍부하여 빅데이터의 적용 및 활용이 가장 기대되는 분야이다. 제조 기업의 공정은 생산설계, 생산, 판매 등의 프로세스가 복잡하게 얽혀있기 때문에 품질 관리와 생산효율성의 증대를 위해 제조 공정 프로세스의 효율화가 중요하다. 본 연구에서는 빅데이터 기술과 프로세스 마이닝 기법을 제조 공정 분석에 접목시킨 빅데이터 클라우드 서비스를 제안한다. 제조 기업은 클라우드 서비스를 활용하여 공정 프로세스의 개선 및 비용절감 등의 효과를 거둘 수 있다. 빅데이터 클라우드 서비스는 공정 프로세스 분석, 공정 시간 분석 등의 다양한 분석 서비스를 제공하며 구현 완료하였다. 사례 연구를 통해 클라우드 서비스의 유효성을 검증하였다.

■ 중심어 : 빅데이터, 제조업, 프로세스 마이닝, 공정 분석

Abstract

Big data is an emerging issue as large data which was impossible to be processed in the past is possible to be handled with the development of information and communication technology. Manufacturing is the most promising field that big data is applied such that there are abundant data available. It is important to improve an efficiency of manufacturing process for quality control and production efficiency because the processes from production design, sales, productions and so on are mixed intricately. This study proposes big data cloud service for manufacturing analysis using a big data technology and a process mining technique. It is expected for manufacturing corporations to improve a manufacturing process and reduced the cost by applying the proposed service. The service provides various analyses including manufacturing analysis and manufacturing duration analysis. Big data cloud service has been implemented and it has been validated by conducting a case study.

■ Keyword : Big data, Manufacturing, Process Mining, Manufacturing Analysis

I. 서론

최근 정보통신기술의 발달로 하루에도 다양한 형태의 수많은 데이터가 생산되고 저장되고 있다. 사건의 발생에 따라 생성되는 데이터는 사실적이며, 적시적소에 활용됨으로써 자산의 효율성 및 효과를 높이고, 비용을 절감하고, 위험을 줄이며, 혁신의 기회를 창출할 수 있다. 과거에는 기술 수준이 부족하여 축적되어있는 방대한 양의 데이터를 활용하지 못하였으나, 데이터 처리 기술이 고도화됨에 따라 점차 데이터의 활용성이

증가하고 있다. 이와 더불어 빅데이터(Big Data)에 대한 관심이 점차 높아지고 있다. 빅데이터는 마케팅, 판매, 생산 등 다양한 분야에서 활용되고 있다. 그 중에서도 제조업 분야는 빅데이터의 활용의 잠재력이 가장 큰 분야이다. 제조업은 산업 초기에 정보시스템을 도입하여 IT 및 자동화 설비 기반이 매우 발달되어 있으며, 생산관리, 품질관리 등에서 데이터를 대량으로 축적하여 방대한 양의 데이터를 보유하고 있다. 현재 제조 산업에서는 빅데이터 분석 활용에 대한 관심이 급증하고 있지만 실제적인 적용이 미미한 실정이다. 제조업에서

빅데이터 분석의 도입에 가장 많은 연구가 수행되고 있는 분야는 프로세스 및 계획 수립 분야이다. 많은 기업들이 제조 공정 프로세스의 효율화를 인식하고 있으며 제조 공정 프로세스 분석을 통해 제품의 품질을 높이고 생산비용을 절감하고자 많은 투자를 하고 있다. 따라서 방대한 양의 제조 산업 데이터를 활용하여 제조 공정 프로세스 분석을 수행하기 위한 기술이 필요하게 되었다.

프로세스 마이닝은 정보 시스템에서 이벤트 로그를 추출 및 분석하여, 업무 수행에 필요한 유용한 정보를 제공하는 연구이다. 실제 작업 프로세스 수행 정보를 바탕으로 공정 프로세스 모델을 도출하고, 도출한 모델을 바탕으로 공정, 작업자 및 설비에 대한 성과 분석을 수행할 수 있다. 이를 통해 업무 수행자에게 객관적인 통찰력을 제공하여, 공정 프로세스의 개선 효과를 기대할 수 있다. 현재 병원, 항만, 관공서 등의 다양한 분야의 비즈니스 프로세스 분석을 위해 많이 활용되고 있지만, 제조 산업의 비즈니스 프로세스 분석은 활발하게 이루어 지지 않고 있다.

프로세스 마이닝을 활용한 제조 공정 프로세스 분석은 공정 데이터를 바탕으로 보다 정확하고 신속한 분석을 가능하게 한다. 이벤트 로그로부터 간결하고 빠르게 공정 프로세스 모델을 도출하고, 도출한 모델을 바탕으로 시간 및 빈도수 관점의 성과 분석, 병목점(Bottleneck Point) 도출 및 공정 유형 분석 등이 가능하다. 뿐만 아니라 공정, 설비 및 작업자, IoT에 대한 상세한 성과 분석을 통하여, 현재의 제조 공정 프로세스의 문제점을 진단하고 개선 방향을 제시할 수 있다.

본 연구에서는 프로세스 마이닝과 빅데이터 기술을 활용하여 제조 분야에 특화된 클라우드 기반의 프로세스 분석 서비스를 제안하고자 한다. 프로세스 마이닝 기법을 적용하여 제조 분야에 특화된 빅데이터 클라우드 분석 서비스를 제공함으로써 산업 현장에서 협업 프로세스 또는 생산 공정이 어떻게 수행되는 가에 대한 현황을 시각화 하여 볼 수 있으며 이를 통해 복잡한 기업 환경에서 조직 및 자원 관리의 투명성을 높이고 경영 효율성을 증대시킬 수 있다. 제안된 서비스는 웹 기반의 클라우드 서비스로 제공하고 있다. 연구에서 제안한 빅데이터 클라우드 서비스는 사례연구를 통해 유효성을 검증하였다.

이후의 구성은 다음과 같다. 제Ⅱ장에서는 제조업에서 빅데이터의 활용과 프로세스 마이닝과 관련된 연구

를 소개하며, 제Ⅲ장은 빅데이터 클라우드 서비스 구성도를 제시한다. 제Ⅳ장에서는 제조 분석 서비스 모듈을 통해 제조 공정 분석을 위해 어떤 기능들이 있는지 설명한다. 제Ⅴ장에서는 구현 결과를 소개하고, 제Ⅵ장에서 사례 연구를 통해 빅데이터 클라우드 서비스를 실증한다. 마지막으로, 제Ⅶ장에서는 결론 및 추후 연구를 소개한다.

Ⅱ. 관련 연구

2.1 제조 산업에서 빅데이터 활용

제조 산업은 많은 선진국들의 근간이 되는 산업으로 산업 혁명 이후 현재에 이르기까지 지속적으로 발전해 온 산업이다. 제조 산업은 컴퓨터 시대가 도래한 이래로 제품의 디자인, 제조, 유통을 위해서 타 산업에 비하여 정보시스템이 초기에 도입하였다. 제품의 품질과 제조 공정의 효율성을 증대시키기 위하여 IT 및 자동화 기반 시설이 탄탄하게 구축되어 있어 양적인 측면과 질적인 측면 모두에서 우수한 데이터를 축적하고 있는 분야이다[11]. 제조업은 2010년에만 약 2 엑사바이트의 새로운 데이터가 발생 및 저장되었고, 그 어떤 분야보다 방대한 양의 데이터를 축적하고 있다[16]. 이러한 데이터는 생산 기계류부터 SCM(Supply Chain Management), 제품 모니터링 시스템에 이르기까지 다양한 소스로부터 발생된다. 최근 제조업의 세계화, 급격한 기술의 발전, 짧아진 제품 수명 주기 등은 글로벌 제조 기업들의 생존을 위협하고 있다. 제조업 전반에서는 이러한 문제들을 해결하고 미래 발전 원동력으로 삼기 위한 해결책으로 빅데이터 기술에 높은 관심과 함께 많은 투자를 하고 있다.

빅데이터는 의사결정을 질을 높이고, 통찰을 발견하고 프로세스의 최적화를 위해 새로운 형태의 처리 방식을 요구하는 대용량, 빠른 속도, 다양한 정보로 이루어져 있는 데이터 집합체를 의미한다[13]. 이러한 빅데이터의 특성은 용량(Volume), 속도(Velocity), 다양성(Variety)의 3V로 대표된다. 최근에는 데이터를 가치로 전환하기 위한 특정한 기술과 분석 방법론을 요구하는 대용량, 빠른 속도, 다양성의 특성을 가진 정보 자산으로 빅데이터의 의미가 더욱 확장되어 받아들여지고 있다[8]. 대용량과 다양성의 빅데이터는 정보기술 분야와 비즈니스 업계에서 기회와 이슈를 동시에 가져오고 있

다[7, 12]. 이슈적인 측면에서는 데이터 관리자들에 대해 빅데이터의 수집, 저장, 처리 등이 주요한 문제로 나타나고 있다[27]. 기획 측면에서는 빅데이터에 기반을 둔 예측과 귀납적인 추론이 비즈니스 분야에서 중요한 역할을 할 것으로 보인다[18].

빅데이터는 제조업에서 R&D 및 설계, 공급사슬, 생산, 판매 및 마케팅, 서비스 등 여러 가치사슬 단계에서 활용 가능성이 존재한다[14]. R&D 및 설계 단계에서 제조 기업들은 빅데이터를 활용하여 제품생산주기를 효과적으로 관리 할 수 있다. 기업들은 제품생산주기관리를 위해 여러 가지 정보시스템들을 도입한다. 도입된 여러 정보시스템들을 통해 수집한 데이터 및 정보들을 하나의 데이터 집합체로 통합하여 관리함으로써 제품생산주기관리의 효율성과 효과를 증대시킬 수 있다. 그리고 판매, 고객 피드백, CRM(Customer Relationship Management) 등의 다양한 데이터를 활용하여 제품 설계 단계에서 좀 더 가치 있는 제품 생산을 가능케 한다. 공급사슬 단계에서 제조 기업들은 판매업체들의 판매 관련 데이터와 재고 데이터 등을 통합적으로 분석하여 수요를 예측하고 재고관리에 반영하여 비용 절감과 효율성 증대, 고객 만족 등의 긍정적인 효과를 높일 수 있다. 생산 단계에서는 사물인터넷(IoT, Internet of Things), RFID 센서 기술 등을 이용하여 각 생산 설비들로부터 실시간으로 생산과 관련된 대량의 데이터를 수집하고 이를 분석함으로써 제조 품질 설비에 대한 예측 정비 실현 등에 기여 할 수 있다. 마지막으로 마케팅 및 판매, 애프터 서비스 단계에서는 고객과의 의사소통에 대한 데이터를 활용하여 실시간으로 제품의 결함 관리, 적절한 고객 컴플레인에 대한 대응 및 생산 프로세스 조정이 가능하다[1]. 이러한 빅데이터 분석을 통해 설계와 생산 단계에서는 각각 50% 비용 절감 효과를 낼 수 있고 7%의 수익 증가 효과를 낼 수 있을 것으로 전망했다[14].

Peter는 제조업에서 빅데이터에 대한 연구 유형 조사 결과, 대부분의 연구가 이론적이며 특정한 문제에 관하여 개념적인 방향에 편중되어 있다고 지적하였다[19]. 다른 연구 유형을 살펴보면 특정한 문제에 대한 솔루션을 제안하거나 사례를 통해 솔루션의 유효성을 검증하는 연구가 다수 존재하였다. 그러나 실제로 구현이 완료되지는 않았지만 새로운 기술에 대한 연구 또는 방법론의 제시와 증명 및 평가를 통해 장단점까지 분석하여 향후 연구 방향을 제시하는 연구는 소수에 불과한 것으로 나타났다. 세부적으로는 제조업에서 빅데

이터에 대한 연구가 애널리틱스를 함께 다루고 있는지를 살펴보았을 때, 52.31%의 연구가 애널리틱스에 대한 언급 없이 오로지 빅데이터 기술에만 관련하여 서술하고 있었으며, 나머지 47.69%의 연구가 어떠한 형태로든 애널리틱스를 빅데이터와 함께 연구주제로 다루고 있었다. 빅데이터 애널리틱스에 다루지는 가장 많은 문제의 유형은 일정 수준의 신뢰도로 의사결정에 도움을 주는 예측의 정확성이다. 예측 분석은 통계학과 같은 다른 분야에서 많은 연구가 이루어져 예측에 관한 방법론과 이론이 풍부하게 있었기 때문에 더욱 활발히 활용되고 있는 것으로 보인다. 그러나 기술, 모델링, 최적화 등의 전문성이 필요한 문제에 관련해서는 제조 산업에서 빅데이터는 유아기 단계로 판단된다.

2.2 제조 산업에서 프로세스 마이닝 활용

프로세스 마이닝은 ERP, MES 등의 다양한 정보 시스템에서 수집한 이벤트 로그를 기반으로 프로세스 관점의 유용한 정보를 제공하는 연구이다[3, 4, 25]. 실제 프로세스 수행 정보를 바탕으로 공정 프로세스 모델을 도출하고, 도출한 모델을 바탕으로 공정, 작업자 및 설비에 대한 성과 분석을 수행할 수 있다. 프로세스 모델 이외에도 의사결정 룰, 소셜 네트워크, 시뮬레이션 모델 등이 도출될 수 있다[20, 21, 23, 27]. 이를 통해 업무 수행자에게 객관적인 통찰력을 제공할 수 있으며, 공정 프로세스 개선에 이바지 할 수 있다. 프로세스 마이닝 기법은 현재 헬스케어, 항만, 관공서 등의 다양한 분야의 비즈니스 프로세스 분석을 위해 많이 활용되고 있다.

제조 산업에서 적용되는 데이터 기술에 관하여 많은 연구가 수행되었는데, 이러한 연구들이 기여한 제조업 분야를 프로세스와 계획 수립, 기업, 유지보수 및 문제 진단, 공급사슬, 유통 및 물류, 제품 설계, 품질의 7가지로 나누어 살펴보았다[19]. 선행연구의 30%가 넘는 수가 프로세스와 계획 수립 분야에서 나타나서 제조업 분야 중에서는 가장 활발하게 연구가 되고 있는 분야로 나타났다. 또한, 프로세스와 계획 수립 분야에서 툴, 방법론, 모델, 아키텍처, 프레임워크, 플랫폼, 이론 등의 다양한 결과물이 나타나고 있다. 제조 산업에서 많은 연구가 프로세스 부문에 집중되고 있는 것은 제조 산업에서 프로세스의 중요성에 대한 근거이기도 하다. 특히 지속적인 최적화를 통해 효율성과 효과가 뛰어난

제조 프로세스는 시장에서 성공적인 비즈니스를 하기 위한 필수적인 전제조건이다[22]. 제조 기업들의 지속적인 프로세스 개선을 위해서 판매, 마케팅 등과 같은 운영 데이터와 제조 프로세스 데이터를 통합하여 분석함으로써 생산과 관련된 전반적인 상황을 파악할 수 있도록 해야 한다[9]. 이에 따라 많은 기업들이 제조 공정 프로세스의 효율화를 인식하고 있으며 제조 공정 프로세스 분석을 통해 제품의 품질을 높이고 생산비용을 절감하고자 많은 투자를 하고 있다. 제조 프로세스 분석을 위한 방법론, 모델, 툴 등이 다수 개발되고 있지만 효율성에서 부족한 것으로 평가 받고 있다[10].

프로세스 마이닝을 활용한 제조 공정 프로세스 분석은 공정 데이터를 바탕으로 보다 정확하고 신속한 분석을 가능하게 한다. 무엇보다, 이벤트 로그로부터 쉽고 빠르게 공정 프로세스 모델을 도출하고, 도출한 모델을 바탕으로 시간 성과 분석, 병목점 도출 및 공정 유형 분석 등이 가능하다. 뿐만 아니라 공정, 설비 및 작업자, lot에 대한 상세한 성과 분석을 통하여, 현재의 제조 공정 프로세스의 문제점을 진단하고 개선 방향을 제시할 수 있다[2].

또한, 제조 공정 데이터는 일련의 분석을 통해 프로세스로 표현될지라도, 프로세스는 다시 하나의 데이터 형태로 변환되어 MES에 투입이 되고, MES는 이 데이터를 기반으로 제조 공정을 관리하게 된다. 이러한 과정을 통해 프로세스(데이터)를 잘 표현하는 것이 가장 중요하며, 이는 데이터의 시각화가 제조 공정 분석에서 가장 중요한 요소 중 하나라는 것을 유추할 수 있다[9]. O'brien은 그의 연구에서 최근 제조기업들은 정보시스템의 발달로 방대한 양의 데이터를 수집 및 저장하고 있으나, 이를 활용하지 못하고 정보 부족 현상을 겪고 있다고 언급하였다. 대용량의 MES 수집 데이터를 각 경영 계층에서 활용할 수 있도록 데이터를 가시화하는 데이터 시각화(Visualization)의 기술 개발을 강조하였다[17]. 선행 연구에서는 분석을 통해 데이터에 기반을 둔 통찰을 제공하고 확실하게 가치 전달을 해야 하는 것에 대한 필요성이 계속해서 나타나고 있다.

III. 빅데이터 클라우드 서비스 구성도

본 연구는 제조 공정 프로세스 분석에 특화된 빅데이터 클라우드 서비스를 제공함으로써 제조 기업의 공

정 프로세스의 개선 및 생산 비용 절감에 기여하고자 한다. 빅데이터 기술의 적용은 대용량의 데이터 처리를 가능케 하며, 클라우드 기술을 통해 분석 서비스에 대한 접근성을 크게 높인다. 프로세스 마이닝을 활용한 분석 서비스를 통하여 공정 프로세스 분석, 공정 시간 분석, 설비 분석, 통과 수율 분석 등의 다양한 분석이 가능하다.

본 연구에서 제안하는 빅데이터 클라우드 서비스는 표준화, 독립성, 확장성을 기본 원칙으로 하여 구축되었다. 서비스 아키텍처는 기존 하둡 등의 오픈 분석 플랫폼 상에서의 업무 데이터를 분석하는 SW를 개발하는 것을 주 목적으로 하였다. 뿐만 아니라 프로세스 분석 모듈은 특정 하둡 구성과 상관없이 일반적인 표준 하둡 아키텍처 상에서 동작할 수 있도록 구성하여 특정 벤더에 대한 종속성 배제함으로써 표준화하였다. 또한, 시스템의 데이터 수집모듈, 데이터 저장소, 프로세스 마이닝 등 각 주요 모듈은 의존성을 줄이고 독립적으로 설계되었다. 이러한 독립적인 모듈은 통합된 뷰를 통해서 일관성 있는 통합 서비스를 가능케 한다. 기존 개발되거나 구성된 대용량 데이터 저장 인프라 및 플랫폼을 최대한 활용할 수 있도록 대중적이면서 호환성이 높은 소프트웨어를 도입하였다. 이는 빅데이터 클라우드 서비스가 데이터 허브로서의 역할을 가능케 함으로써 확장성을 확보하였다.

클라우드 기반 제조 프로세스 분석 서비스는 일반적인 IaaS¹⁾ 상에 구축되는 SaaS(Software as a Service)로 필요 시 데이터 및 분석 API(Application Program Interface)를 오픈 할 수 있는 플랫폼으로 확장 가능하다. 본 연구의 제조 공정 분석을 위한 빅데이터 클라우드 서비스의 구성도는 <그림 1>과 같이 크게 하둡 데이터 플랫폼과 SaaS 플랫폼으로 구성된다. 하둡 데이터 플랫폼은 빅데이터 대용량 데이터 처리를 위한 플랫폼으로 기업의 MES(Manufacturing Execution System), ERP(Enterprise Resource Planning) 등의 정보시스템에서 추출된 데이터의 저장 및 프로세스 분석용 데이터로의 변환의 역할을 수행한다. SaaS 플랫폼은 프로세스 마이닝을 적용한 분석 서비스를 제공하는 어플리케이션 중심으로 이루어져 있다.

1) IaaS(Infrastructure as a Service)는 서버, 스토리지, 네트워크를 가상화 환경으로 만들어, 필요에 따라 인프라 자원을 사용할 수 있게 서비스를 제공하는 형태.



〈그림 1〉 빅데이터 기반 제조 프로세스 분석 서비스 구성도

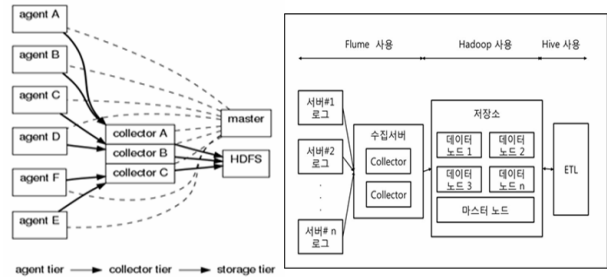
3.1 하둡 데이터 플랫폼

하둡은 맵리듀스와 하둡 분산 파일 시스템(HDFS, Hadoop Distributed File System)로 구성되어 있다. 구글에서 제작한 오픈소스로서 페이스북, 아마존, 트위터 등과 같은 IT 기업들이 사용하고 있는 하둡은 빅데이터와 같은 대용량 데이터 처리의 능력 향상을 위해 전략적으로 개발된 기술이다[5, 24]. 본 연구에서는 하둡의 대중성과 확장성을 고려하여 데이터 플랫폼을 구축하는데 있어서 기반 기술로 채택하였다.

하둡 데이터 플랫폼은 프로세스 데이터 저장장치, 프로세스 로그리더, 프로세스 정보 추출, 프로세스 이벤트 저장장치로 구성된다. 프로세스 데이터 저장장치는 기업의 MES, ERP 등의 정보시스템에서 추출한 데이터를 저장 및 관리하는 기능을 수행한다. 프로세스 로그리더는 다양한 형태의 로그 포맷을 읽을 수 있도록 하며, 머신 및 공정 데이터로부터 프로세스 정보를 추출해 내는 프로세스 정보 추출기가 하둡 데이터 플랫폼에 탑재되어 제조 현장의 로그 데이터를 이벤트 리포지토리 기능을 수행하는 프로세스 이벤트 저장장치에 저장한다.

하둡 데이터 플랫폼을 구성하는 하둡 머신은 IaaS 플랫폼 상의 가상 서버에 위치하여 데이터의 분석 처리 용량과 사이즈에 따라 추가적인 노드 확장이 가능하도록 구성되어 지속적인 확장이 가능한 구조로 이루어져 있다. 제조 프로세스 분석 서비스 모듈과 분산 저장소 간의 인터페이스를 위해 플럼(Flume)을 중심으로 스콥(Sqoop), HBase 등의 하둡 에코 시스템들을 유기적으로 결합하여 최신 하둡 기술들이 접목된 분산 저장소를 구축하였다. 플럼은 Agent, Collector, Repository로 이뤄지는 3-Tier 구조를 가지며, 에이전트(Agent) 노드는 로그를 생산하는 시스템에 설치되어 플럼의 컬렉트 티

어로 데이터를 전송하고 이는 다시 하둡 저장소에 데이터를 저장하게 함으로써 데이터 수집 모듈의 유연성, 확장성, 안정성을 확보하였다. 플럼의 구조는 <그림 2>와 같다. 기업의 데이터 보안을 보장하기 위하여 데이터 저장 및 전송 시 암호화 모듈을 적용하였다.



〈그림 2〉 플럼(Flume) 구조

각 모듈간 버전 종속성을 파악 및 제조 프로세스 분석 서비스 모듈과의 인터페이스 및 병렬 분산 처리의 성능에 영향을 미치는 설정 항목들을 도출하여 호튼웍스, 클라우드라 등 하둡 배포 환경을 벤치마킹하고 필요 시 통합 배포판을 활용하여 유지 보수 및 개발이 용이한 데이터 플랫폼을 구성하였다. 기본적으로 모든 머신 로그 데이터는 이벤트 리포지토리에 표준화된 형태로 저장되며 인메모리(In-Memory) 기반 쿼리 및 분석 처리를 통하여 기존 하둡이 가지는 맵 리듀스 배치 작업의 응답 속도 저하 문제를 해결하였다.

3.2 SaaS 플랫폼

SaaS 플랫폼은 하둡 데이터 플랫폼에서 수집 및 저장된 데이터를 이용하여 현장 사용자 및 데이터 분석가가 쉽게 사용할 수 있는 통합 분석 시스템을 클라우드 소프트웨어 형태로 서비스를 제공한다. 사용자가 직접 데이터를 업로드하고 분석까지 수행할 수 있을 뿐만 아니라 데이터를 가공하여 재다운로드 할 수 있는 기능까지 제공하고 있다. 또한 클라우드 기반의 서비스는 사용자의 접근성을 높여서 인터넷을 이용하여 시간과 장소에 제한 없이 이용이 가능하다. 접근성 및 사용자의 데이터 이용 및 분석 자율성을 확보하여 사용자 측면에서 편의성을 극대화하였다.

제조 프로세스 분석 서비스는 데이터 수집에서부터 공정 분석, 이후 개선점 도출과 프로세스 모니터링까지 제공하는 서비스이다. 우선, 다양한 시스템으로부터 나

오는 시스템 로그를 기반으로 생산 제조 현장의 공정 프로세스 패턴 및 문제점과 이에 대한 개선점 도출한다. 도출된 개선 방안은 현장에 반영되어 제조 공정 프로세스를 최적화한다. <그림 3>과 같이 제조 공정 분석 클라우드 서비스는 데이터 수집부터 분석에 이르기까지 제조 공정 분석을 광범위하게 포괄하며 이 모든 것을 서비스로 제공한다. 이러한 공정 프로세스 분석은 일회성에 그치지 않고, 최적화된 프로세스를 기반으로 개선 전 후 비교 분석 등을 포함한 반복적인 분석을 수행함으로써 지속적으로 공정 프로세스를 모니터링하고 생산성 및 품질 개선을 위한 데이터 수집 및 분석 도구로 활용된다.



<그림 3> 제조 공정 분석 클라우드 서비스(SaaS)

IV. 제조 프로세스 분석 서비스 모듈

제조 프로세스 분석 서비스 모듈은 데이터 플랫폼과 SaaS 플랫폼 양측 내부에서 독립적으로 작동하도록 구성되어 있다. 프로세스 로그 리더, 프로세스 정보 추출기, 프로세스 데이터 리포지토리, Process Analyzer 네 가지의 모듈로 구성되어 있으며, 제조 프로세스 분석 서비스 모듈의 구성도는 <그림 4>와 같다. Process Analyzer를 제외한 프로세스 로그 리더, 프로세스 정보 추출기, 프로세스 데이터 리포지토리는 앞서 설명하였듯이 하둡 데이터 플랫폼에 포함되어 있으며 각각 데이터의 표준화, 프로세스 분석에 필요한 정보 추출, 데이터 저장의 역할을 수행한다. Process Analyzer는 프로세스 분석을 수행하고 결과를 제공하는 분석 서비스의 핵심 역할을 담당한다. 각각의 모듈은 서로 상호작용하여 제조 공정 프로세스 분석을 수행하도록 최적화 되어 있다.

프로세스 로그 리더는 데이터 수집과 데이터 표준화 기능을 수행한다. MES, ERP의 정보시스템뿐만 아니라 기존의 레거시 시스템으로부터 정형 및 비정형 데이터

를 수집한다. 그리고 데이터를 프로세스 분석이 가능하도록 표준화한다. 여기서 표준화는 프로세스 마이닝의 적용이 가능하도록 케이스, 이벤트, 시간, 작업자의 네 가지 정보를 설정하고 표준 이벤트 로그 형태로 데이터를 변환하는 것이다. 데이터 표준화 과정에서 프로세스 로그 리더가 데이터 칼럼 헤더를 자동적으로 인식하여 케이스, 이벤트, 시간, 작업자의 네 가지 정보를 기본적으로 설정하며, 사용자가 직접 칼럼 헤더를 매칭할 수도 있다.



<그림 4> 제조 프로세스 분석 서비스 모듈

프로세스 정보 추출기는 수집된 데이터의 표준화 과정에서 발생하는 여분의 속성 데이터를 활용하여 프로세스 정보를 추출한다. 각 정보는 어떤 데이터를 사용하는냐에 따라서 공정 정보에서부터 센서 정보까지 다양한 정보가 될 수 있으며 제조 공정 분석에 활용된다.

프로세스 데이터 리포지토리는 하둡 분산 파일 시스템을 기반으로 표준화 과정을 통해 분석 대상 데이터가 변환된 표준 이벤트 로그를 저장한다. 또한, Process Analyzer에서 분석 중에 발생하는 가공 데이터를 저장하는 저장소 역할을 수행한다. 데이터의 보안을 위해서 데이터를 중심으로 정밀한 감사 및 제어를 실시하여, 누가, 언제, 어디서 데이터를 열람하는지 상시 모니터링하고, GUI(Graphic User Interface) 도구로 다수의 계정에 대한 데이터 접근 권한을 조정하여 데이터를 관리한다.

Process Analyzer는 분석 결과를 막대 그래프, 원 그래프, 간트 차트, Day-hour 차트 등 다양한 방식으로 시각화한다. 분석 시각화 UI(User Interface)는 HTML5 기반으로 서비스함으로서 특정 브라우저에 대한 종속성을 줄이고, 반응형 웹 디자인(Responsive Web Design)을 채택하여 태블릿 등 다양한 스크린 크기를 가지는

디바이스에 대해서도 일관성 있는 화면을 제공한다. 프로세스 분석을 위해서 사용되는 프로세스 패턴 데이터, 이벤트 및 케이스 관점 등의 다차원 관점의 빈도수 및 시간 등에 대한 통계 결과 등을 저장 및 관리하는 기능을 수행한다. 프로세스 분석 중에 필요에 따라서는 프로세스 분석을 통해 한 번 가공된 데이터를 익스포트하여 R, 시뮬레이션 등의 외부 도구와 연계하여 추가적인 분석이 가능하다. 또한, 복수의 기관 및 사용자가 개별로 접속하여 독립적인 데이터 집합을 관리할 수 있도록 구축되어 있어서 이러한 기능들은 Process Analyzer의 데이터 분석 기능뿐만 아니라 일종의 데이터 허브 역할까지 수행토록 한다.

V. 서비스 구현-Process Analyzer

본 연구에서는 제조 공정 분석을 위한 빅데이터 클라우드 서비스를 구현하였다. 앞서 설명하였듯이 연구를 통해 개발된 시스템은 클라우드 형태로 서비스를 제공하며 웹 기반이므로 장소에 따른 제약이 없이 공정 분석에 특화된 다양한 분석 도구를 제공한다. 공정 프로세스 분석, 공정 시간 분석, 설비 분석, 통과수율 분석, 부하 분석의 제조 공정 분석이 가능하며 세부 기능으로는 데이터 등록 및 관리, 대시보드, 데이터 검색, 프로세스 맵 도출, 애니메이션, 프로세스 패턴 분석, 성과분석, 선후행 분석, 선후행 비교 분석, 수율 분석, 부하 분석 및 지연 분석, 소셜 네트워크 분석 등이 있다. 구현 화면은 <그림 5>와 같다.



<그림 5> 구현 화면

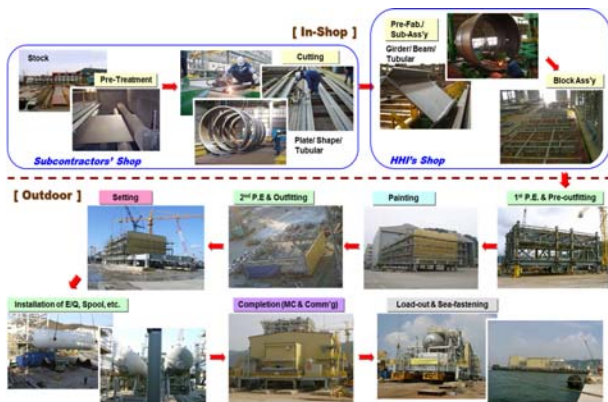
대시보드는 케이스 수, 이벤트 수, 작업자 수, 데이터 수집 기간 등의 데이터 개요를 나타낸다. 프로세스 맵 기능은 프로세스 마이닝 알고리즘을 적용하여 케이스,

이벤트, 시간, 작업자의 간단한 정보만으로 프로세스 맵을 도출한다. 애니메이션 기능은 도출된 프로세스 맵에 기반하여 애니메이션을 작동시킨다. 애니메이션을 통해 프로세스에서 나타나는 정체구간을 직관적으로 시각화하며 세부 옵션으로 특정 시간대의 데이터만을 애니메이션으로 볼 수 있도록 설정 가능하다. 프로세스 패턴 분석 기능은 데이터 내에서 나타나는 프로세스 패턴을 기준으로 모든 케이스들을 그룹화하여 어떠한 패턴이 가장 많이 나타나는지, 소요시간이 오래 걸리는지 등을 확인 할 수 있으며, 각 패턴을 Dotted Chart를 이용하여 시각화하여 볼 수 있다. 성과 분석은 시간 및 빈도수 등의 성과 수치를 통계적으로 산출하고, 각 케이스, 이벤트, 작업자, 부서 별로 수치를 세분화할 수 있다. 선후행 분석은 작업 간의 선후행 관계를 통해 기존의 조직 내의 작업 순서가 얼마나 잘 준수되고 있는지 볼 수 있으며, 작업 간의 밀접도를 빈도수를 기준으로 확인한다. 또한, 선후행 비교 분석을 통해서 다른 데이터와 선후행 관계를 비교해 봄으로써 프로세스의 표준화에 기여한다. 부하 분석 및 공정 지연 분석을 통해 공정 지연이 발생하는 원인을 찾아낼 수 있다. 부하 분석은 기준 처리 능력을 넘어서는 과도한 작업량을 발견하고 어떤 작업자 혹은 부서에서 발생하는 지에 관하여 조사하여 개선방안을 도출할 수 있도록 한다. 공정 지연 분석은 계획과 실적 데이터를 비교하여 각 작업들이 스케줄에 맞게 진행되고 있는지를 파악하여 향후 스케줄에 반영할 수 있도록 한다. 이러한 분석들은 공정 프로세스의 개선 및 관리를 효율화하고 공정 지연을 줄이는 효과를 가져 올 수 있다.

VI. 사례 연구

개발된 시스템의 유효성을 검증하기 위하여 실제 데이터에 분석 서비스를 적용한 사례연구를 수행하였다. 데이터를 제공한 기업은 A 조선소로 조선사업, 해양플랜트 사업을 주로 하고 있다. 본 사례연구에서는 해양플랜트 공정 데이터를 바탕으로 빅데이터 클라우드 서비스를 적용해보았다. A 조선소가 제공한 데이터는 해양플랜트 제조 프로젝트 관리 및 일정 데이터이다. 해양플랜트 공정 프로세스는 <그림 6>과 같이 원재료들을 가공하여 도장, 절단, 페인팅, 설치, 조립 등의 일련의 작업을 거치면서 블록으로 제작하고, 블록을 다시 모듈화 및

조립하여 해양플랜트 설비를 제조하게 되는 과정으로 이루어져 있다. 대부분의 원자재 가공은 외주업체를 통해 작업이 이루어진다. 데이터는 계획과 실적 두 가지의 정보를 포함하고 있다. 계획 데이터는 A 조선소에서 프로젝트 수행 계획 정보를 포함하고 있으며, 실적 데이터는 현재까지의 프로젝트 진행사항에 대한 실제 발생한 작업 및 일정에 관한 정보를 담고 있는 데이터이다.

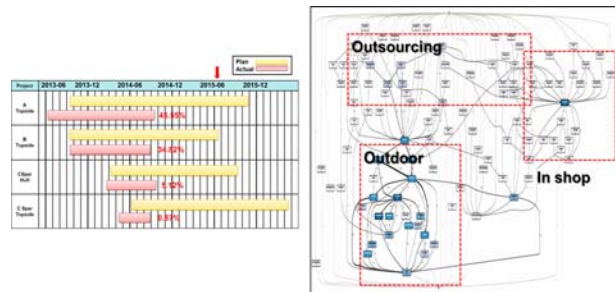


〈그림 6〉 해양플랜트 공정 프로세스

해양플랜트 사업은 많은 인력과 자원이 투입되는 규모가 매우 큰 프로젝트 기반 사업이다. 사업 특성상 이러한 다수의 프로젝트가 동시다발적으로 진행되기 때문에 해양플랜트 공정 프로세스의 관리가 어렵고, 공정 지연의 문제가 빈번하게 발생한다. A 조선소도 해양플랜트 프로젝트를 관리하는데 많은 어려움을 겪고 있었다. 특히 다수의 프로젝트 동시 진행하고 있어, 여러 프로젝트에서 공정 지연 문제를 찾는 것에 큰 어려움을 겪고, 공정의 지연이 빈번하게 일어났다. 기존의 방법으로 공정 지연의 원인 파악하는데, 매우 오랜 시간이 소요 되었는데, 원인 파악을 위해 엑셀을 이용한 데이터 처리, 통계적 기법, 인터뷰 등을 이용하고 있었다. 또한, 프로젝트 공정 관리 측면에서는 다수의 프로젝트를 각각 단일적으로 관리하고 있는 정보시스템이 존재 하였으나, 다수의 프로젝트들이 동시에 진행되면서 조선소 내부의 작업 흐름이 매우 복잡해지게 되는데 이를 관리하는 데 있어서 어려움이 존재하였다. 복잡한 해양 플랜트 공정 프로세스의 가시화와 공정 지연의 문제를 해소하기 위하여 빅데이터 클라우드 서비스를 적용하였다.

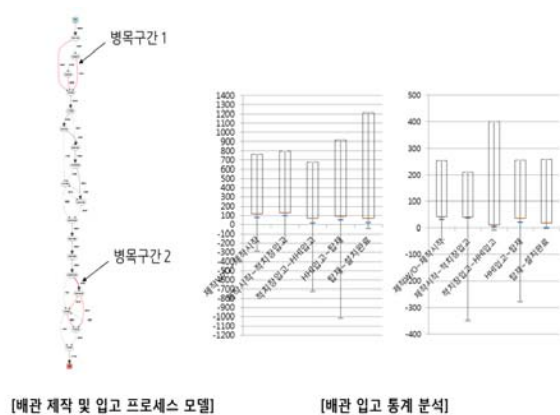
Process Analyzer를 바탕으로 A 조선소의 해양플랜트 제조 공정 데이터를 분석하였다. 우선, 다수의 프로젝트

가 동시다발적으로 진행됨으로써 발생하는 전체적인 프로젝트 관리 문제를 해결하기 위해 복수의 프로젝트를 고려해서 분석을 진행하였다. <그림 7>과 같이 4개의 프로젝트에서 계획, 실적 데이터를 비교하여 작업 현황을 파악할 수 있었다. 4개의 프로젝트로부터 생성된 데이터를 바탕으로 <그림 7>과 같이 공정 프로세스를 도출하여 전체적인 작업흐름을 시각화 하고, 프로세스 모델에서 작업들을 외주(Outsourcing), 야드 및 야외(Outdoor), 공장 내부(In Shop) 세 가지의 범주로 특정 지을 수 있었다.



〈그림 7〉 공정 프로세스 모델 도출

도출된 프로세스 모델에 기반한 성과분석을 통한 작업처리시간에 따른 병목구간을 파악하여 1차적인 공정 지연의 문제를 진단할 수 있다. 이를 활용하여 배관제작 및 설치 프로세스 데이터를 분석하여 병목구간을 발견하고, 결과를 바탕으로 배관입고에 대한 통계분석을 수행하였다.



〈그림 8〉 배관 제작 및 입고 프로세스 분석

빅데이터 클라우드 서비스를 적용하여 분석 결과 몇 가지 의미 있는 결과를 얻을 수 있었다. 첫 번째로 원인 파악의 시간 단축이다. 기존의 기업에서는 공정 지연의 원인을 파악하기 위하여 엑셀을 이용한 데이터 처리 및

통계적 기법을 사용하였으나 이는 수개월의 시간이 소요되었으며, 파악된 원인도 수개월 전 데이터를 기반으로 한 것이었기 때문에 현실적으로 개선하기 어려웠다. 하지만 빅데이터 클라우드 서비스의 적용을 통해 단시간 내에 문제점을 발견하여 즉각적인 개선과제 도출이 가능하게 되었다. 두 번째로 실질적인 해양플랜트 공정 프로세스를 시각화하여 작업흐름을 파악하고, 동시에 다발적인 해양 플랜트 설비 프로젝트의 관리가 가능하게 되었다. 해당 기업의 MRP(Manufacturing Resource Planning) 시스템은 각각의 프로젝트 스케줄 및 제조 과정을 단일적으로 관리하였다. 그러나 실제로는 다수의 프로젝트가 동시에 진행되면서 제조 공정이 섞이게 되어 전체적인 시점에서의 공정 프로세스 관리가 필요하였다. 해양 플랜트 공정 데이터를 기반으로 전체적인 관점의 새로운 프로세스 모델을 도출하고 공정 지연의 원인을 파악할 수 있었다. 또한, 프로젝트 관리를 위한 계획된 일정 데이터와 실질적으로 진행되고 있는 실적 데이터를 비교하여 어느 부분에서 작업의 지연과 선수행이 발생하는지 파악하여 향후 스케줄을 탄력적으로 조정할 수 있도록 하였다. 세 번째로 병목구간 파악을 통한 작업흐름의 개선이다. 시간 관점의 공정 지연의 문제뿐만 아니라 작업량에 관한 병목점을 발견하여 작업 흐름의 개선에 도움을 줄 수 있었다. 특정 부서에서 작업량이 과도하게 많이 발생하는 것을 파악하여 작업 일정을 조정하고 작업량을 분배하는 조치를 통해 병목현상을 해소하였다.

VII. 결론 및 추후 연구

본 연구에서는 프로세스 마이닝 기법과 빅데이터 기술을 활용하여 제조 분야에 특화된 클라우드 기반의 공정 분석 서비스를 제안하고 이를 구현하였다. 하둡 데이터 플랫폼에 기반하여 빅데이터 대용량처리가 가능한 클라우드 서비스는 확장성, 접근성, 사용자 편의성을 고려하여 구현되었다. 제조 산업에서 빅데이터에 관한 연구는 아직 유아기 단계로 판단되는 상황에서 프로세스 마이닝 기법을 접목한 빅데이터 클라우드 형태의 분석 서비스를 개발한 것은 유례가 없는 연구로 제조 산업을 비롯한 프로세스 마이닝, 빅데이터의 연구 분야에 새로운 시도이다. 또한 제조 산업에서 빅데이터 연구 유형이 이론적이며 개념적인 방향에 편중되어 있는 상황에서 시스템 구현을 완료하고, 사례 연구를 수행하였다. 또한,

제조 공정 데이터에 맞추어 직관적인 통찰을 제공하기 위해 다양한 차트 및 그래프를 사용하여 시각적 효과를 높였다. 제조 산업의 급변하는 외부 환경과 치열한 경쟁 속에서 앞으로는 보다 객관적인 데이터에 기반하여 문제를 신속하고 정확하게 진단하여 비용절감과 경영 효율화를 위한 지속적인 개선 및 최적화가 필요하다. 이러한 관점에서 본 연구에서 구현한 빅데이터 클라우드 서비스는 제조 기업들의 지속적인 공정 프로세스 개선 및 모니터링에 도움을 줄 수 있다.

향후 연구로, 실시간 분석에 대한 연구가 필요하고, 예측 관점에서 성과 분석을 통해 도출된 작업자 수, 작업시간 등의 성과지표를 기반으로 프로세스 및 스케줄에 대한 시뮬레이션을 통해 공정 프로세스와 향후 스케줄의 최적화에 대한 연구가 필요하다.

참고 문헌

- [1] 노규성, 박상휘, “제조실행 시스템에의 빅데이터 적용방안에 대한 탐색적 연구”, 한국디지털정책학회, 2014.
- [2] 손속영, Bernardo Nugroho Yahya, 송민석 외 다수, “프로세스 마이닝 기반 공정 프로세스 분석: 삼성전기 사례 연구”, 한국CAD/CAM학회 동계학술대회논문집, pp.739-744, 2014.
- [3] 송민석, 김덕영, 류광열, 조용주, 조현제, “프로세스 마이닝 활용한 c-MES 공정 프로세스 분석 방안”, 한국생산제조 시스템학회 학술발표대회논문집, pp. 193-193, 2011.
- [4] 송민석, 정일교, 조용주, 조현제, “프로세스 마이닝을 활용한 생산 공정 데이터 분석”, 대한산업공학회 춘계학술대회논문집, pp.326-345, 2012.
- [5] Aisling, O., D. Jurate, and D.S. Roy, “‘Big data’, Hadoop and cloud computing in genomics”, *Journal of Biomedical Informatics*, Vol.46, No.2, pp.774-781, 2013.
- [6] Chen, H., R. Chiang, and V. Storey, “Business Intelligence and Analytics: From Big Data to Big Impact”, *MIS Quarterly*, 2012.
- [7] Chen, M., S. Mao, and Y. Liu, “Big data: a Survey”, *Mobile Networks and Applications*, Vol.19, No.2, pp. 171-209, 2014.

- [8] De Mauro, A., M. Greco, and M. Grimaldi, "What is big data? A consensual definition and a review of key research topics", *AIP Conference Proceedings*, pp. 97-104, 2015.
- [9] Groger, C., F. Niedermann, and B. Mitschang, "Data Mining-driven Manufacturing Process Optimization", *Proceedings of the World Congress on Engineering*, Vol.3, 2012.
- [10] Hernandez-Matias, J.C., A. Vizan, A. Hidalgo, and J. Rios, "Evaluation of techniques for manufacturing process analysis", *Journal of Intelligent Manufacturing*, Vol.17, No.5, pp.571-583, 2006.
- [11] Hessman, T., "Putting Big Data to Work", *Industryweek*, pp.14-18, 2013.
- [12] Labrinidis, A. and H.V. Jagadish, "Challenges and opportunities with big data", *Proceedings of the VLDB Endowment*, Vol.5, No.12, pp.2032-2033, 2012.
- [13] Laney, D., "The Importance of 'Big Data': A Definition", *Gartner*, 2012.
- [14] McKinsey Global Institute, "Big data: The next frontier for innovation, competition, and productivity", 2011.
- [15] Milagros, R. and M. Ernesto, "Agent Learning in Autonomic Manufacturing Execution Systems for Enterprise Networking", *Computers and Industrial Engineering*, Vol.64, pp.901-925, 2012.
- [16] Nedelcu, B., "About Big Data and its Challenges and Benefits in Manufacturing", *Database Systems Journal*, Vol.4, No.3, pp.10-19, 2013.
- [17] O'Brein, D., "Make sure MES delivers data-driven value", *Plant Engineering*, Vol.7, pp.53-56, 2011.
- [18] Ohlhorst, F., "Big data analytics: turning big data into big money", *John Wiley and Sons*, 2013.
- [19] Peter, O. et al., "Big data in manufacturing: a systematic mapping study", *Journal of Big Data*, 2015.
- [20] Rozinat, A., M.T. Wynn, W.M.P. van der Aalst, A.H.M. Hofstede, and C.J. Fidge, "Workflow simulation for operational decision support", *Data and Knowledge Engineering*, Vol.68, No.9, 2008, Sixth International Conference on Business Process Management (BPM)-Five selected and extended papers, pp.834-850, 2009.
- [21] Rozinat, A., R.S. Mans, M. Song, and W.M.P. van der Aalst, "Discovering simulation models", *Information Systems*, Vol.34, No.3, pp.305-327, 2009.
- [22] Slack, N., S. Chambers, and R. Johnston, "Operations Management", 6th ed. Harlow: Financial Times Prentice Hall, 2010.
- [23] Song, M. and W.M.P. van der Aalst, "Towards comprehensive support for organizational mining", *Decision Support Systems*, Vol.46, No.1, pp.300-317, 2008.
- [24] Taylor, R.C., "An overview of the Hadoop/MapReduce /HBase framework and its current applications in bioinformatics", *BMC Bioinform*, Vol.11, 2010.
- [25] van der Aalst, W.M.P., "Process Mining: Discovery, Conformance and Enhancement of Business Processes", Springer, Heidelberg, Germany, 2011.
- [26] van der Aalst, W.M.P., M.H. Schonenberg, and M. Song, "Time prediction based on process mining", *Information Systems*, Vol.36, No.2, Special Issue: Semantic Integration of Data, Multimedia, and Services, pp.450-475, 2011.
- [27] Wu, X., X. Zhu, G.Q. Wu, and W. Ding, "Data mining with big data", *Knowledge and Data Engineering IEEE Transactions on*, Vol.26, No.1, pp.97-107, 2014.

저자 소개



이 용 혁(Yong-Hyeok Lee)

- 2012년 : 울산과학기술대학교 테크노경영학부 (학사)
- 2012년~현재 : 울산과학기술원 경영공학과 (공학석사)
- 관심분야 : 프로세스 마이닝, 데이터 분석



송 민 석(Min-Seok Song)

- 2006년 : 포항공과대학교 산업경영공학과 (공학박사)
- 2006년~2009년 : 아인트호벤 공과대학교 (박사 후 과정)
- 2010년~현재 : 울산과학기술원 경영학부 (교수)
- 관심분야 : 프로세스 마이닝, BPM(Business Process Management), 비즈니스 분석(Business Analytics)



하 승 진(Seung-Jin Ha)

- 1990년 : 인하대학교 산업공학과 (공학사)
- 1995년 : 인하대학교 대학원 산업공학과 (공학석사)
- 1995년~현재 : 현대중공업
- 관심분야 : 제조 빅데이터 활용, 제조시스템 개선(Industry 4.0 등) TPS (Toyota Production System) 적용, 제조 Simulation 응용



손 숙 영(Sook-Young Son)

- 2013년 : 울산과학기술원 경영학과 (학사)
- 2014년 : 울산과학기술원 경영공학과 (공학석사)
- 2014년~현재 : 현대중공업 공정지원부 (대리)
- 관심분야 : 프로세스 마이닝, 통계분석, 빅데이터



백 태 현(Tae-Hyun Baek)

- 1993년 : 울산대학교 산업공학과 (공학사)
- 1995년 : 울산대학교 산업공학과 (공학석사)
- 1995년~현재 : 현대중공업 공정지원부 (부장)

- 관심분야 : IoT/ICT, Smart Factory, 프로세스 마이닝