

# 조선 산업에서 프로세스 마이닝을 이용한 블록 조립 프로세스의 계획 및 실적 비교 분석

## Comparison between Planned and Actual Data of Block Assembly Process using Process Mining in Shipyards

이동하(Dongha Lee)\*, 박재훈(Jae Hun Park)\*\*, 배혜림(Hyerim Bae)\*\*\*

### 초 록

본 논문에서는 조선 산업에서 블록 조립 작업에 대한 계획 프로세스와 실적 프로세스를 비교하는 방법을 제안한다. 제안한 방법은 계획과 실적 데이터 기반으로 프로세스 마이닝 기법을 이용하여 프로세스 모델을 도출하고 비교 분석을 수행하는데, 분석 절차는 1) 데이터 전처리, 2) 분석 수준의 정의, 3) 조립 블록 군집화, 4) 군집별 프로세스 모델 도출, 5) 계획과 실적 프로세스 모델 비교 다섯 단계로 구성된다. 단계 5에서는 프로세스 모델, 작업, 프로세스 인스턴스, 모델 적합도와 같은 다양한 관점에서 계획과 실적의 프로세스를 비교할 것을 제안하고, 각 관점별 비교 인자들을 정의한다. 특히, 적합도 관점에서는 교차 적합도를 정의하여, 도출된 프로세스 모델에 대해 자신의 데이터에 대한 적합도뿐만 아니라, 상대 데이터(계획 모델의 경우 실적 데이터, 실적 모델의 경우 계획 데이터)에 대한 적합도를 계산하고 비교 분석할 것을 제안한다. 제안한 방법의 효용성은 세계 최고 수준의 국내 조선업체의 블록 조립 계획 시스템과 블록 조립 모니터링 시스템의 실제 데이터를 이용하여 사례 연구를 통해 설명하고 검증하였다.

### ABSTRACT

This paper proposes a method to compare planned processes with actual processes of block assembly operations in shipbuilding industry. Process models can be discovered using the process mining techniques both for planned and actual log data. The comparison between planned and actual process is focused in this paper. The analysis procedure consists of five steps : 1) data pre-processing, 2) definition of analysis level, 3) clustering of assembly blocks, 4) discovery of process model per cluster, and 5) comparison between planned and actual processes per cluster. In step 5, it is proposed to compare those processes by the several perspectives such as process model, task, process instance and fitness. For each perspective, we also defined comparison factors. Especially, in the fitness perspective, cross fitness is proposed and analyzed by the quantity of fitness between the discovered process model by own data and the other data(for example, the fitness of planned model to actual data, and the fitness of actual model to planned data). The effectiveness of the proposed methods was verified in a case study using planned data of block assembly planning system (BAPS) and actual data generated from block assembly monitoring system (BAMS) of a top ranked shipbuilding company in Korea.

**키워드** : 조선, 블록조립, 계획데이터, 실적데이터, 프로세스마이닝, 프로세스모델, 적합도  
Shipbuilding, Block Assembly, Planned Data, Actual Data, Process Mining, Process Model, Fitness

본 논문은 2012년도 교육과학기술부의 재원으로 한국연구재단의 기초연구 사업(No. 2012R1A1A2008335)과 (주) 대우조선해양의 지원을 받아 수행된 것입니다.

\* Central R&D Institute, Daewoo Shipbuilding and Marine Engineering Co., Ltd.(dongha@dsme.co.kr)

\*\* Department of Industrial Engineering, Pusan National University(pjh3479@pusan.ac.kr)

\*\*\* Corresponding Author, Department of Industrial Engineering, Pusan National University (hrbae@pusan.ac.kr)

2013년 08월 22일 접수, 2013년 09월 03일 심사완료 후 2013년 9월 27일 게재확정.

## 1. 서 론

조선 산업에서 생산을 자동화 하고 작업을 효과적으로 관리하고자하는 노력은 지속되어 왔다. 특히 국내 조선 산업의 경우 세계 선두 자리에 있으나 중국의 추격으로 인해 보다 확고한 경쟁우위를 확보하기 위한 지속적인 기술 개발 및 효과적인 작업 관리가 절실하다. 이러한 바램과 더불어 국내 조선 사업은 IT 기술의 발달과 함께 다양한 생산 공정 모니터링 시스템을 구축하여 적용하고 있다. 생산 공정 모니터링 시스템은 보다 효과적인 생산 관리를 가능하게 하여 국내 조선 산업의 기술력과 경쟁력을 높이는데 기여할 수 있다. 특히, 최근에는 전세계적으로 빅 데이터(big data)에 대한 관심 및 수요가 증가함에 따라 조선 산업에서도 축적된 데이터들을 활용한 작업 분석 및 성능 개선에 관심을 가지기 시작 하였다.

본 논문은 선박의 많은 생산 공정 중에서 가장 기본이 되고 복잡한 조립 블록의 작업 공정에 대해서 다룬다. 선박은 블록 건조 공법으로 생산 되는데, 이는 거대한 선박을 작업장 내에서 조립할 수 있는 적절한 크기의 블록으로 나누고, 이들 블록들을 작업장 내에서 조립한 후, 외장, 도장 작업 등을 거쳐 도크(dock)에서 최종 탑재하여 선박을 건조하는 방식이다[3, 4]. 이때 조립 작업장 내에서 조립되는 단위 블록들을 조립 블록(assembly block)이라고 한다. 조립 블록은 작업장 내에서 작업할 수 있는 수준의 크기로 분할되지만 가로, 세로의 길이가 수 미터에 이르고, 무게가 수십 톤에 달하며, 많은 수의 조립품들이 여러

조립 작업 단계를 거쳐서 완성된다. 또한 하나의 조립 블록은 선박의 일부분이므로 구조적 특성상 좌/우현의 유사블록이 존재할 뿐 동일한 블록이 없고 이를 이루는 조립품의 구성도 모두 다른 특징을 가진다. 따라서 생산 계획자는 각 조립 블록에 대해서 매번 작업 단계를 정의하고 생산계획을 수립해야 하는 어려움이 있다. 또한 이렇게 생산해야 하는 블록의 다양성뿐만 아니라, 조선 산업의 특성상 작업장이 넓고 협력업체 등 많은 작업 조직과 작업자에 의해서 생산이 이루어지기 때문에 실제 생산 공정을 모두 이해하고 파악하는 것이 힘들다. 뿐만 아니라 정확한 생산 계획과 관리를 어렵게 하는 요인이 된다. 이러한 이유로 조선 산업에서는 IT 기술의 발달과 함께 지속적으로 보다 나은 생산 공정 모니터링 시스템 개발에 많은 노력을 기울여 왔고 최근에는 안정화 되어 많은 데이터가 축적되고 있는 실정이다.

본 논문에서는 이러한 조선 산업의 요구사항에 대해서, 최근 연구가 활발하게 수행되고 있는 프로세스 마이닝 기법을 이용하여 생산을 이해하고 분석할 것을 제안한다. 생산 공정 모니터링 시스템에 축적된 이벤트 로그 데이터를 이용하여 프로세스 모델을 도출하고, 생산 계획자가 계획 단계에서 파악하고 있는 프로세스와 비교, 개선해 나가면 생산관리 기술을 향상시킬 수 있다. 또한 조선 산업에서 지속적으로 고민해 온 계획과 실적의 비교 부분에 있어서도 제안하는 방법이 좋은 대안이 될 수 있다.

본 논문에서는 블록 조립 계획 시스템으로부터 계획 데이터를, 블록 조립 모니터링 시

시스템으로부터 실적 데이터를 추출해서 각각 프로세스 모델을 도출하고 비교하는 방안을 제안한다. 제안한 방법의 타당성은 실제 세계 선두 그룹의 국내 조선소의 데이터를 이용하여 검증하였다.

## 2. 프로세스 마이닝과 관련 연구

### 2.1 프로세스 마이닝

프로세스 마이닝은 기업 정보시스템에서 프로세스 실행 과정에서 발생한 이벤트 로그로부터 의미 있는 지식을 추출해 내는 과정이다. 즉, 프로세스의 액티비티 실행 과정에서 누적된 기록을 통하여 프로세스의 개선이나 설계에 필요한 유용한 지식을 추출하는 것이다. 프로세스 저장소의 실행 결과 및 이벤트에 관한 로그를 추출하여 기존의 통계 기법[7], 인공지능 기법[11, 12], 사회적 네트워크 기법[2] 등을 이용하여 분석함으로써 프로세스 모델을 유도하거나, 업무의 상관 관계, 작업자의 업무 전달 관계 등 프로세스 수행 상의 특징을 분석한다[1].

프로세스 마이닝에는 세 가지 기본 유형 및 단계가 있다. 첫째 프로세스 도출(process discovery), 둘째 프로세스 모델의 적합도 검사(conformance checking), 셋째 프로세스 모델의 개선 및 향상(enhancement)이다. 프로세스 도출은 사전 정보 없이 정보시스템의 이벤트 로그들로부터 프로세스 모델을 생성하는 기법을 말한다. 이는 프로세스 마이닝에서 가장 중요한 시작 단계에 해당된다. 적합도 검사는

로그에 기록된 현실이 모델에 일치하는지 또는 모델이 현실에 일치하는지를 확인 하는 것인데, 도출된 프로세스 모델을 평가하는 척도로 이용 된다. 프로세스 모델의 개선 및 향상은 이벤트 로그에 기록된 실제 프로세스 관련 정보를 이용하여 기존의 프로세스 모델을 확장하고 개선하는 것이다. 적합도 검사는 모델과 현실 사이의 정합성을 측정하는 반면에, 모델의 개선 및 향상은 이전 모델을 바꾸거나 확장시키는 것을 목표로 한다. 프로세스 마이닝 기술의 활용 관점에서는 이 세 가지 유형이 모두 단계적으로 이루어졌을 때 큰 의미를 가진다고 하겠다. 따라서 본 연구에서도 실적 데이터 기반의 생산 프로세스 도출에 머무는 것이 아니라, 계획과 실적 프로세스의 효율적인 비교 방안까지 제안하고자 한다.

프로세스 마이닝의 출발점은 이벤트 로그이다. 프로세스 마이닝 기법은 이벤트가 발생한 순서대로 기록되어 있다는 것을 가정한다. 하나의 이벤트는 프로세스의 단위 작업을 나타내며, 하나의 케이스와 관련을 맺고 있다. 이벤트 로그는 또한 이벤트에 대한 부가적인 정보를 저장하고 있을 수 있는데, 이를 활용해서 추가적인 분석이 가능하다. 데이터 마이닝 기법을 이용해서 프로세스를 분석하는 연구[6]도 있는데, 프로세스 모델을 발견하고 모델 기반의 다양한 분석을 수행하고자 한다면 프로세스 마이닝 기법의 활용이 더욱 효과적이다. 최근 조선 산업의 경우 IT 융합 기술의 발달과 함께 공정 모니터링 시스템이 개발되고 안정화 되어서 프로세스 마이닝 기법의 적용을 위한 많은 데이터가 축적되어 있는 것이 또한 본 연구의 기대 효과를 높이고 있다.

## 2.2 관련 연구

프로세스 마이닝에 대한 초기 연구는 주로 이벤트 로그들로 부터 프로세스 모델을 추출하는 내용이 주를 이루었다. 그리고 주로 의료 및 서비스 분야와 같이 파악하기 힘든 복잡한 프로세스를 이벤트 로그를 이용해서 발견하고 분석 하는데 적용되었다. 프로세스 모델 추출을 위한 대표적인 마이닝 기법은 휴리스틱 마이닝(heuristic mining), 퍼지 마이닝(fuzzy mining), 알파 마이닝( $\alpha$  mining) 등이 있다. 휴리스틱 마이닝은 이벤트간 프로세스의 흐름과 빈도를 이용해서 프로세스 모델을 도출하는 방법이고, 퍼지 마이닝은 복잡한 프로세스를 프로세스의 발생 빈도를 기준으로 주요 프로세스와 상대적으로 덜 중요한 프로세스를 구분할 수 있는 알고리즘을 제공한다. 알파 마이닝 알고리즘은 이벤트 집합 내의 프로세스 패턴을 확인함으로써 페트리넷(Petri Net)을 도출할 수 있는데, 이는 적합도 검사에 유리한 특징을 가진다.

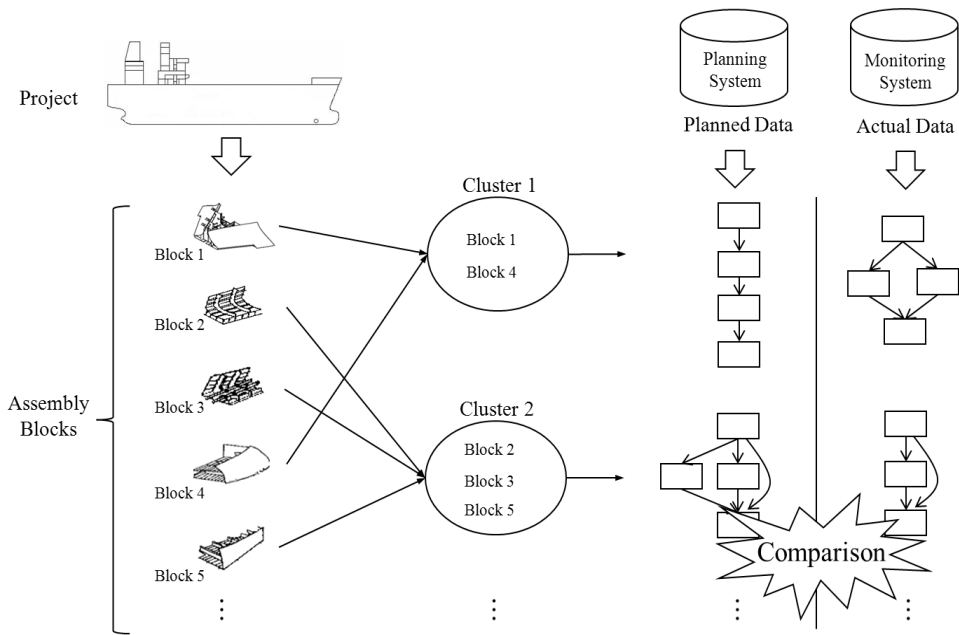
프로세스 도출을 위한 다양한 마이닝 기법들이 지속적으로 개발되고 있을 뿐만 아니라, 적용 분야도 최근에는 금융[14], 통신[5], 제조 산업 등 다양한 영역으로 확대되고 있다. 조선 산업의 경우도 최근에 마이닝 기술을 이용한 트랜스포터의 블록 이동 프로세스 분석에 대한 연구 사례가 있다. 데이터 마이닝 기법을 이용해서 트랜스포터의 블록 이동 특성에 따라 블록을 군집화 하고, 각 군집별로 프로세스 모델을 분석한 연구가 있고[10], 프로세스 마이닝 기법을 이용해서 이벤트 로그로부터 전체 프로세스를 이해하고 예외 프로세스를 감지하기까지의 분석 프레임워크를 제안한

연구도 있다[9].

그러나 프로세스 마이닝을 이용한 블록 조립 작업 분석에 대한 연구는 찾아보기 힘들다. 특히 본 논문에서는 프로세스 마이닝 기법을 이용하여 계획과 실적 프로세스의 비교에 초점을 두고 있는 점이 기존 연구와 두드러진 차이점 이다. 블록 이동 프로세스 분석에 대한 연구는 작업장에서 조립 블록이 제작되고 나서 탑재 되기까지 트랜스포터를 이용한 블록의 이동 패턴을 찾고, 예외 이벤트와 프로세스를 발견하는 것이 목적이지만, 블록 조립 작업은 조립 블록이 완성되기까지 많은 작업 단계들이 계획과 동일한 순서로 지정한 작업장에서 작업이 수행되었는지에 관심을 가진다. 따라서 효율적인 계획과 실적의 비교가 요구된다.

## 3. 제안 분석방법

본 논문에서 제안하는 분석 방법에 대한 개념은 <Figure 1>과 같다. 각기 다른 구조적 특성을 가지는 조립 블록들을 작업 특성을 고려하여 군집화 하고, 각 군집별 프로세스 모델을 도출한 후 계획과 실적의 작업 특성을 비교하는 것이다. 군집 내에 존재하는 블록들 간에도 블록의 구조나 작업 단계가 완전히 동일한 것은 아니기 때문에 프로세스 마이닝 기법을 이용해서 프로세스 모델을 도출할 것을 제안한다. 이렇게 도출된 모델은 해당 군집의 작업 프로세스의 특징을 표현하게 된다. 이를 비교하면 각 조립 블록별로 계획과 실적을 하나씩 비교하는 것보다 쉽고 빠르게 작업 특징을 잘 파악할 수 있다. 실제 생산 계획자들도

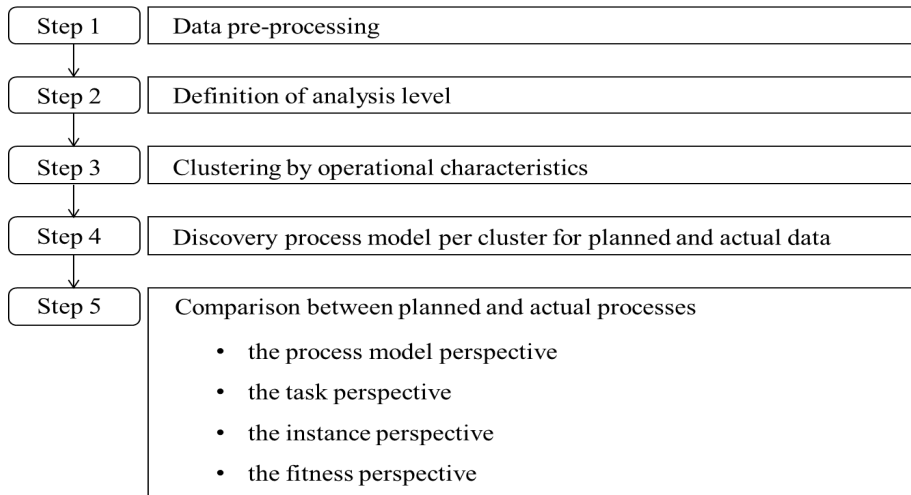


<Figure 1> Concept of Analysis Method

작업의 유형별로 ‘계획 케이스’ 데이터를 정의하고 활용하고 있는데 해당 블록이 어느 ‘계획 케이스’에 포함되는지를 판단해서 동일한 작업 공기와 작업장 활용을 결정하게 된다. 따라서 계획자들이 이미 정의한 ‘계획 케이스’ 기준으로 블록을 군집화 하고 계획과 실적 데이터를 비교해 보면, 계획자들이 경험에 의해서 정의하고 판단한 ‘계획 케이스’별 속성 정보대로 실제 생산 현장에서 작업이 이루어지는지 분석할 수 있다. 그리고 ‘계획 케이스’에는 작업 단계별 공기와 작업장 정보를 가지고 있을 뿐, 작업 프로세스에 대한 대표 모델을 가지고 있지는 못하다. 이는 모든 조립 블록의 구조적 특성이 차이가 나기 때문인데, 프로세스 마이닝 기법을 활용하면 유사한 작업 특성의 ‘계획 케이스’ 즉 군집에 대해 프로세스 모델을 쉽게 생성할 수 있다. 본 연구에서

는 계획자들이 경험에 의해 정의해 둔 ‘계획 케이스’를 이용해서 정의한 블록의 계획 속성대로 군집화를 하고 계획 데이터 기반의 모델과 실적 데이터 기반의 모델을 비교하고자 한다.

제안한 분석 방법을 프로세스 마이닝 관점에서 단계적으로 표현하면 <Figure 2>와 같이 다섯 단계로 정의할 수 있다. 단계 1은 분석 데이터를 얻기 위한 데이터 전처리 과정이고, 단계 2에서는 데이터 분석 수준을 결정한다. 분석 목적에 따라서 어떤 데이터를 분석 기준의 액티비티(activity)로 선정할지를 결정해야 한다. 정보시스템에 존재하는 이벤트로 그들은 분석을 위해서 만들어진 데이터가 아니기 때문에 효율적인 분석을 위해서 추가적으로 필요한 정보는 무엇인지, 그리고 어떤 데이터를 기준으로 분석을 수행할지 결정해야 한다. 단계 3에서는 각기 다른 작업 특징을 가



〈Figure 2〉 Analysis Procedures

지는 조립 블록들을 작업 특성에 따라서 군집화 한다. 본 연구에서는 계획자들이 계획 단계에서 이용한 ‘계획 케이스’ 데이터 기준으로 조립 블록을 군집화 한다. 생산 계획자가 관리하는 계획 기준과 실적을 비교하기 위해서이다. 단계 3에서 정의한 군집을 기준으로 단계 4에서는 동일 군집별로 계획 데이터와 실적 데이터 기준으로 각각 프로세스 모델을 도출한다. 이때 활용되는 프로세스 마이닝 기법 및 조건은 동일하게 설정한다. 결국 동일한 기준으로 프로세스 모델을 도출하고 비교하는 것이다. 끝으로 단계 5에서는 도출된 프로세스 모델을 다양한 관점에서 비교한다. 본 논문에서는 네 가지 관점에서의 분석을 제안하고 각 관점별 비교 인자들을 정의한 후 분석한다. 이러한 분석 방법은 조선 산업과 같이 넓은 작업장에서 다양한 작업을 가지고 작업 프로세스의 정확한 표준을 설정하기 힘든 경우에, 프로세스 모델을 찾고 계획과 실적을 비교하는데 유용할 것으로 사료된다. 각 관점별 분석 기준

인자들은 사례 연구에서 자세히 설명한다.

## 4. 사례 연구

조선 산업의 블록 조립 계획 시스템으로부터 계획 데이터를 추출하고, 블록 조립 공정 모니터링 시스템으로부터 실적 데이터를 추출하여 제안한 방법으로 사례연구를 수행하였다. 동일한 선종 및 선형의 두 개의 프로젝트를 대상으로 전처리 과정을 거쳐 프로세스 모델을 도출하고 분석을 수행하였다. 각 단계별 주요 내용은 다음 각 절에서 자세히 설명된다.

### 4.1 데이터 전처리(Step 1)

데이터 마이닝 및 프로세스 마이닝 분야에서는 데이터 전처리 과정이 매우 중요하다. 전처리 결과에 의해서 분석 데이터의 품질이 달라지고 분석 결과에 많은 영향을 주기 때문

이다. 본 사례 연구에서는 다음 세 가지 경우를 고려해서 데이터 전처리를 수행하였다.

#### (1) 사외 제작 조립 블록 데이터 제외

사외 제작 조립 블록 데이터의 경우, 계획 관점에서는 데이터가 존재하지만, 실제로 작업이 이루어지는 과정에 대한 실적 데이터는 모니터링 되지 않기 때문에 분석 대상에서 제외한다. 결국 사내에서 제작되는 조립 블록에 대해서만 계획과 실적 데이터를 비교하는 것이다. 물론 이후 사외에서 제작되는 조립 블록도 사내 수준과 동일하게 실적 데이터가 모니터링 된다면 동일한 방법으로 분석이 가능하겠다.

#### (2) 최종 조립 블록이 완성되기까지의 데이터가 없는 블록 정보는 제외

최종 조립블록이 완성되기까지의 프로세스 모델을 도출하고 비교하기 때문에, 실적 데이터 관점에서 최종 조립블록까지 완성되지 않은 블록 데이터는 제외한다. 각 조립 블록에 대해서 최종 대조립 작업 실적 데이터가 존재하는지 확인하고 존재하지 않는 경우, 해당 블록의 데이터를 분석 대상에서 제외한다.

#### (3) 재작업, 수정작업 등과 같이 추가된 작업 단계 데이터 제외

실제 생산현장에서 작업된 모니터링 시스템의 실적 데이터를 확인해 보면 작업 후 재작업이나 수정작업이 이루어진 내용도 데이터를 통해서 확인할 수 있다. 그러나 본 연구에서는 조립 블록의 각 작업에 대해서 실제 계획한 내용과 실적 프로세스를 비교하는 것을 목적으로 하기 때문에 추가 작업에 대한 데이터는 해당 분야의 전문가 의견에 의해 분석 대상에서 제외 한다. 이는 조립 블록 자체를 제외하는 (2)의 경우와 달리, 블록의 해당 작

업 단계만 제외한다. 별도로 이들 재작업과 수정작업의 유형 및 특성에 대해서도 분석을 수행할 수 있지만 실적 기반의 실제 작업과 계획 단계에서 정의한 작업 프로세스와의 차이 분석을 통해 계획 단계의 표준 프로세스를 개선하고 활용하는 것을 목표로 하기 때문에 대상에서 제외한다.

프로세스 분석 관점의 데이터 단위로 언급하면 (1)과 (2)의 데이터 전처리는 프로세스 인스턴스를 제외하는 것이고, (3)의 경우는 이벤트를 제외시키는 경우에 해당된다. 이렇게 데이터 전처리 후에 두 개의 프로젝트에 대해서 135개의 조립 블록 데이터를 획득하였다. 결국 135개의 프로세스 인스턴스를 기준으로 분석을 수행하는 것이다.

## 4.2 분석 수준 정의(Step 2)

선박의 한 부분을 이루는 조립 블록은 많은 작업 단계를 거쳐서 조립된다. 또한 선종과 선형에 따라서도 블록의 구조가 달라지므로 작업 단계가 다를 수 밖에 없다. 즉 동일한 선박을 제작하는 경우를 제외하고는 동일한 조립 블록이 존재하지 않는다. 따라서 조립 블록의 작업을 주요 작업 특성을 기준으로 군집화 한 후, 자원을 할당하고 일정계획을 수립하는 것이 필요하다.

하나의 조립 블록에 대해서 모니터링 시스템에 저장된 작업 단계(work stage)별 데이터에 대한 예는 <Table 1>과 같고, 이를 구조화 하면 <Figure 3>과 같다. 상선의 경우는 20~40개의 작업 단계를 가지고, 해양 프로젝트의 경우는 많게는 60개의 작업 단계로 구성

〈Table 1〉 Event Log of an Assembly Block

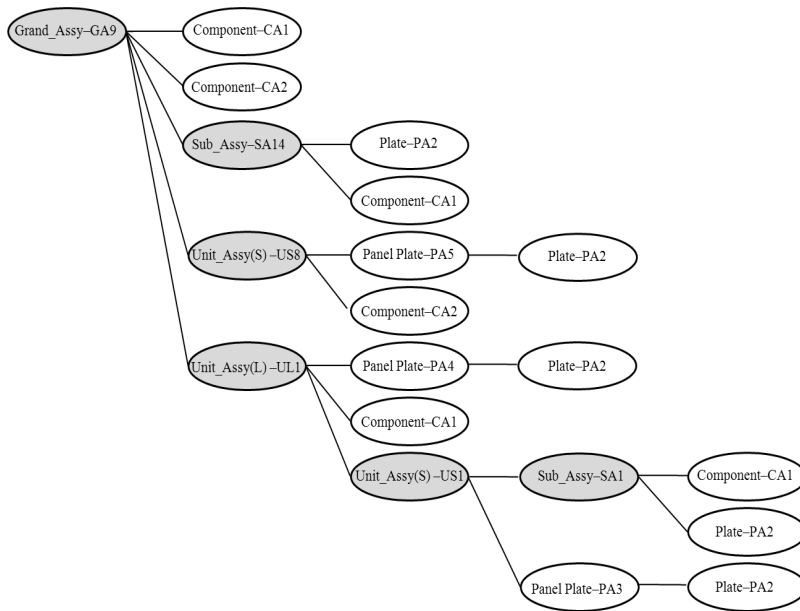
Project	Block	Work Type	Work Stage Code	Work Shop	Bay	Finish Date Time	Organization
1000	Block 1	Component	CA1	Component Shop	1	26 May 09:00	A
1000	Block 1	Plate	PA2	Plate Shop	2	30 May 09:00	B
1000	Block 1	Sub Assembly	SA1	Assembly Shop 3	1	17 May 17:49	C
1000	Block 1	Plate	PA2	Plate Shop	3	03 June 09:00	B
1000	Block 1	Panel Plate	PA3	Plate Shop	3	03 June 09:00	B
1000	Block 1	Unit Assembly (Small)	US1	Assembly Shop 3	1	13 June 13:00	C
1000	Block 1	Component	CA1	Component Shop	2	08 June 09:00	A
1000	Block 1	Component	CA2	Component Shop	2	15 June 09:00	A
1000	Block 1	Plate	PA2	Plate Shop	3	10 June 09:00	B
1000	Block 1	Component	CA1	Component Shop	1	08 June 09:00	A
1000	Block 1	Plate	PA2	Plate Shop	3	13 June 09:00	B
1000	Block 1	Panel Plate	PA5	Plate Shop	3	13 June 09:00	B
1000	Block 1	Panel Plate	PA4	Plate Shop	3	10 June 09:00	B
1000	Block 1	Unit Assembly (Large)	UL1	Assembly Shop 3	3	21 June 06:42	D
1000	Block 1	Plate	PA2	Plate Shop	2	16 June 09:00	B
1000	Block 1	Unit Assembly (Small)	US8	Assembly Shop 3	1	21 June 13:00	C
1000	Block 1	Sub Assembly	SA14	Assembly Shop 3	1	26 June 06:47	C
1000	Block 1	Component	CA2	Component Shop	2	25 June 09:00	A
1000	Block 1	Component	CA1	Component Shop	2	20 June 09:00	A
1000	Block 1	Grand Assembly	GA9	Assembly Shop 3	3	01 July 02:11	D

된다. 따라서 본 연구에서는 주요 작업 유형(work type)에 대해서만 프로세스 모델을 도출하고 분석할 것을 제안한다. <Table 1>과 <Figure 3>에서 주요 작업은 음영으로 표시되었다. 주요 작업 유형은 자재 혹은 중간 조립품들이 결합되는 소조립(Sub Assembly), 중조립(Unit Assembly), 대조립(Grand Assembly) 단계로 정의한다. 그리고 중조립의 경우는 다시 소형(Small Type), 곡형(Curved Type), 대형(Large Type)으로 구분한다. 이는 작업 유형에 따라 작업의 난이도와 자원의 활용이 달

라지기 때문이다. 결국 소조립(Sub\_Assy), 소형 중조립(Unit\_Assy(S)), 곡형 중조립(Unit\_Assy(C)), 대형 중조립(Unit\_Assy(L)), 대조립(Grand\_Assy)의 다섯 가지 작업 유형이 대상이 되고, 각 작업 유형은 이후 설명에서 괄호 안의 표기로 설명된다.

<Table 1>에서 보여지는 작업 단계 코드(work stage code)를 활용하면 작업 유형에 대해서 전체 블록 조립 작업 단계에서 어느 위상에 위치하는지를 확인할 수 있다. <Figure 3>의 예에서 작업 유형이 동일하게 소조립이





<Figure 3> Work Stages of an Assembly Block

지만 작업 단계가 ‘SA1’인 경우는, 해당 소조립이 소형 중조립 단계에서 조립되고, 작업 단계가 ‘SA14’인 경우는 대조립 단계에서 최종 조립된다. 따라서 이벤트 로그 데이터에서 분석 대상을 선정하는 기준은 작업 유형으로 하지만, 이후 작업 분석은 작업 단계 코드를 기준으로 수행한다. 또한 계획과 실적 관점에서 작업장 활용이 어떻게 바뀌었는지도 비교 분석하기 위해서 프로세스 마이닝을 위한 분석 액티비티 선정은 작업 단계 코드와 작업장으로 결정한다. 이는 이후 프로세스 모델에서 ‘작업 단계 코드-작업장’의 형태로 각 노드에 표현된다.

### 4.3 조립 블록 군집화(Step 3)

선박의 한 부분인 각 조립 블록은 물리적 구조 및 작업 단계가 모두 다르므로 동일한

작업 특성을 가지는 조립 블록들에 대해서 군집화가 요구된다. 프로세스 마이닝 이전에 먼저 프로세스 로그를 군집화 하는 방법으로 로그 군집화(log clustering)[8]와 자취 군집화(trace clustering)[13]에 관한 연구가 있다. 그러나 본 연구에서는 계획자가 관리해 온 조립 블록의 작업 특성 구분 기준을 존중하고, 계획과 실적 데이터 기준의 프로세스 비교 분석에 중점을 둔다.

주요 작업 특성 관점에서 조립 블록을 구분하고 자원 할당과 일정 계획을 수립하기 위해서, 생산 계획자는 ‘계획 케이스(Plan Case)’ 데이터를 관리 한다. 특정 유형을 가지는 작업 특성에 대해서 작업 단계별 공기와 작업장을 경험에 의해서 정의해 두고, 계획 수립 시 기준 정보로 활용 한다. 사례 연구 대상 블록들에 대해서 계획 단계에서 선택된 ‘계획 케이스’와 각 케이스별 특징은 <Table 2>와 같다.

〈Table 2〉 Plan Case(Cluster)

Plan Case (PC)	No. of Blocks (Instances)	Characteristics of PC mentioned by domain expert
PC 1	48	<ul style="list-style-type: none"> <li>• Final block is assembled in shop 1</li> <li>• Unit_Assy(L) is manufactured in the line-typed bay of work shop</li> </ul>
PC 2	16	<ul style="list-style-type: none"> <li>• Final block is assembled in shop 2</li> <li>• Duration of Grand_Assy is over 5 days</li> </ul>
PC 3	15	<ul style="list-style-type: none"> <li>• All work stages are performed in shop 4</li> <li>• Shape of assembly block is general</li> </ul>
PC 4	12	<ul style="list-style-type: none"> <li>• All work stages are performed in shop 5</li> <li>• Shape of assembly block is general</li> </ul>
PC 5	8	<ul style="list-style-type: none"> <li>• Final block is assembled in shop 2</li> <li>• Special part is served via shop 3</li> </ul>
PC 6	8	<ul style="list-style-type: none"> <li>• All work stages are performed in shop 3</li> </ul>
PC 7	8	<ul style="list-style-type: none"> <li>• Final block is assembled in shop 1</li> <li>• Unit_Assy(L) is manufactured in the fixed-typed bay of work shop</li> </ul>
PC 8	6	<ul style="list-style-type: none"> <li>• Final block is assembled in shop 5</li> <li>• Welding amount of block is much</li> </ul>
PC 9	4	<ul style="list-style-type: none"> <li>• Final block is assembled in shop 2</li> <li>• Welding amount of block is much</li> </ul>
PC 10	4	<ul style="list-style-type: none"> <li>• All work stages are performed in shop 2</li> </ul>
PC 11	4	<ul style="list-style-type: none"> <li>• Final block is assembled in shop 3</li> <li>• Unit_Assy(C) is manufactured in shop 2</li> </ul>
PC 12	2	<ul style="list-style-type: none"> <li>• Special shaped block is assembled in shop 5</li> </ul>

각 조립 블록에 대해서 계획 수립 시, ‘계획 케이스’를 선택하면 해당 계획 정보가 조립 블록에 반영되고 이를 기준으로 계획을 확정 혹은 일부 수정하게 된다. 조립 블록의 형상 및 작업 단계 구성을 확인하고 가장 적합하다고 판단되는 ‘계획 케이스’를 계획자가 선택하는 것이다. <Table 2>에서 보여지는 것처럼, 사례 연구의 대상 조립 블록에 대해서는 ‘PC 1’의 유형으로 계획 케이스가 가장 많이 선택되었는데, 이는 대형 중조립 작업이 작업장의 흐름 라인에서 작업된 후, 최종 조립 블록이 ‘Shop 1’에서 조립되는 특징을 가진다. 이와 유사하지만 ‘PC 7’의 경우는 대형 중조립 작업이 고정 라인에서 작업이 된다. 또한 ‘PC 8’

과 ‘PC 9’의 경우는 모두 용접 작업량이 많은 조립 블록이지만 ‘Shop 5’와 ‘Shop 2’에서 각각 작업을 수행하는 차이를 보인다. ‘PC 12’의 경우는 곡이 매우 심한 특별한 형상의 조립 블록에 대한 계획 정보를 별도로 정의한 것이다. 동일한 ‘계획 케이스’에 속한다고 판단하는 조립 블록들도 각 작업 단계별로 보면 작업 프로세스가 조금씩 다르기 때문에 계획자가 해당 군집을 대표하는 프로세스 모델을 정의하는 것은 힘들다. 본 연구에서는 계획자에 의해 관리 되고 있는 ‘계획 케이스’를 동일한 작업 특성을 가지는 군집으로 간주하고, 계획과 실적 데이터 기반으로 각각 프로세스 모델을 도출하고 작업 특성을 비교한다.

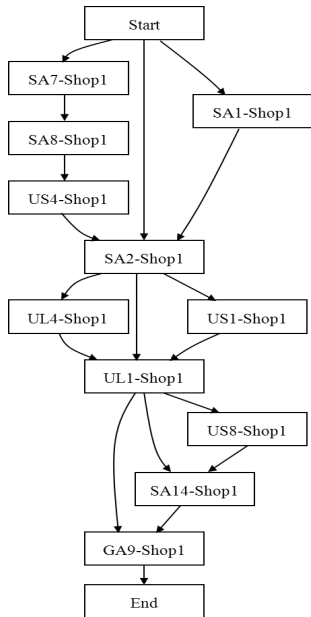
#### 4.4 프로세스 모델 도출(Step 4)

단계 4에서는 계획 시스템의 계획 데이터와 조립 모니터링 시스템의 실적 데이터로부터 ‘계획 케이스’로 구분되는 블록 군집별로 각각 프로세스 모델을 도출한다. 프로세스 모델은 대표적인 프로세스 마이닝 기법인 휴리스틱 마이닝을 적용하였다. 휴리스틱 마이닝은 이벤트의 발생 빈도와 이벤트 간의 작업 순서를 고려해서 프로세스 모델을 도출하는 기법으로 임계값(threshold)을 이용해서 노이즈(noise)와 모델의 불완전성을 제거할 수 있는 특징을 가진다[15]. 따라서 군집에 포함된 다양한 조립 블록에 대해서 임계값 조정을 통해 작업 특성을 분석하기가 용이하다. 단계 3에서 동일한 작업 특성을 가지는 군집으로 활용하기로 한 ‘계획 케이스’ 별로 계획 데이터와 실적 데이터 기반의 프로세스 모델을 각각 도출 후 프로세스를 비교 분석한다.

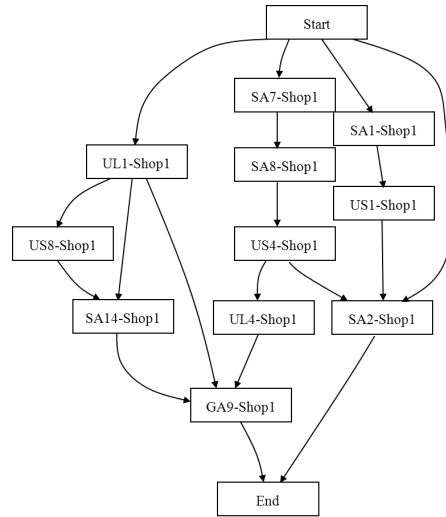
프로세스 마이닝을 위해서는 분석하고자 하는 데이터에 대해서 케이스(case), 액티비티(activity), 타임스탬프(timestamp), 수행 주체(originator) 등을 정의해야 한다. 케이스는 프로세스 인스턴스라고도 하는데 프로세스를 분석하는 단위가 된다. 본 사례 연구에서는 조립 블록의 작업을 분석하는 것이므로 프로젝트의 조립 블록이 케이스가 된다. 액티비티는 작업(task)이라고도 하며 해당 프로세스의 단위 작업을 의미하므로 조립 블록의 각 작업 단계가 액티비티가 된다. 본 사례 연구에서는 제 4.2절에서 설명한 것처럼 작업 단계와 작업장을 액티비티로 정의하여 분석한다. 타임스탬프는 각 작업 단계의 완료일로 선정하였는데, 이는 완료 기준으로 작업의 순서와 일

정이 준수되었는지를 더 중요하게 관리하기 때문이다. 수행 주체는 별도로 선정하지 않고 작업 단계와 작업장의 관점에서 액티비티 기준으로 분석 하였다.

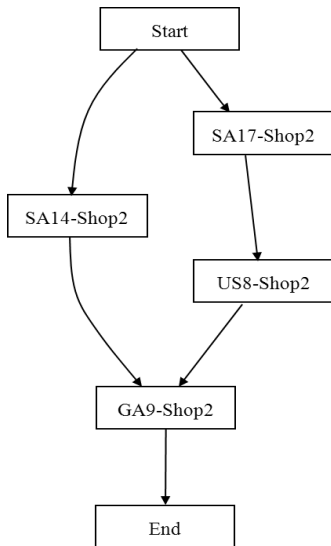
군집 내 가장 블록 수가 많은 ‘PC 1’과 ‘PC 2’에 대해서 계획과 실적 데이터 기반으로 각각 휴리스틱 마이닝을 수행한 결과의 프로세스 모델은 <Figure 4(a), (b)>, <Figure 5(a), (b)>와 같다. 파라미터 설정은 휴리스틱 마이닝의 기본 임계값[15]을 사용하였는데, ‘양의 관측 임계값(positive observations threshold)’은 ‘1’로 설정 했다. 즉, 한번 이상 관측되는 작업 흐름에 대해서도 모델에 반영하여 어떤 특징들이 있는지 우선 비교 분석하기 위함이다. <Figure 4(a), (b)>, <Figure 5(a), (b)> 프로세스 모델을 통해서 먼저 작업장 관점에서는 계획자가 파악하고 있던 대로 ‘PC 1’의 경우는 계획과 실적 모두, 모든 작업 단계가 ‘Shop 1’에서 이루어지고, ‘PC 2’의 경우는 ‘Shop 2’에서 작업이 이루어지는 것을 확인할 수 있다. 프로세스의 흐름 관점에서는 ‘PC 1’의 경우는 <Figure 4(a)>의 계획 모델에서는 ‘UL1’ 작업 단계까지 순차적으로 작업이 이루어지다가, ‘UL1’ 작업 단계 이후 작업이 다양한 순서로 이루어짐을 보인다. 그리고 ‘Start’ 노드에서 ‘SA2’ 작업 단계로의 연결선이 있다는 것은 ‘SA2’가 가장 먼저 작업된 프로세스 인스턴스가 존재한다는 것을 의미한다. <Figure 4(b)>의 실적 모델에서는 계획 모델과 달리, ‘UL1’의 작업부터 가장 먼저 작업된 경우가 있음을 보인다. 이는 소조립 작업의 완료 없이 먼저 중조립 작업이 수행된 경우에 해당 되므로 해당 프로세스의 인스턴스를 찾아서 분석할 필요가 있다. 또한 <Figure 4(b)> 와 <Figure



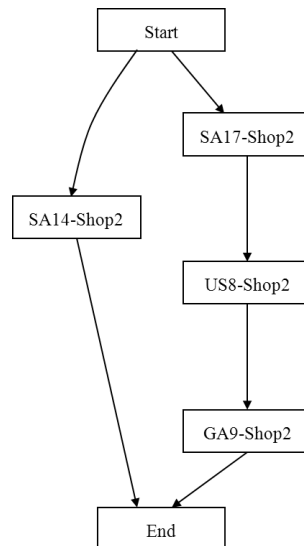
〈Figure 4(a)〉 Plan Model of PC 1



〈Figure 4(b)〉 Actual Model of PC 1



〈Figure 5(a)〉 Plan Model of PC 2



〈Figure 5(b)〉 Actual Model of PC 2

5(b)>의 실적 프로세스 모델에서 ‘SA2’에서 ‘End’로 ‘SA14’에서 ‘End’로 이어지는 연결선이 각각 존재하는 것을 볼 수 있는데, 이는 항상 ‘GA9’ 작업 단계에서 최종 작업을 완료한다는 계획과 달리, 일부 간단한 소조립의 경우는 대조립 완료 후에도 작업이 이루어지고 있음을 나타낸다.

물론 ‘양의 관측 임계값’과 작업 간의 ‘의존 임계값(dependency threshold)’을 높이면 보다 일반화된 프로세스 모델을 획득할 수 있다. 그러나 우선 이렇게 낮은 임계값으로 군집 내 프로세스 인스턴스의 작업 특성을 확인해 보면 예상하지 못했던 실적 결과를 계획 데이터 기반의 모델과 비교해 볼 수 있다. 다만 이는 전체 프로세스의 대표성 보다는 부분적으로 발생하는 예상치 못한 흐름이나, 군집 내 프로세스 인스턴스 간의 작업 흐름 차이를 발견하는데 유리하다고 하겠다.

#### 4.5 계획과 실적간 프로세스 비교

본 논문은 조선 산업의 블록 조립 작업 특성상 블록 유형별 프로세스 모델을 정의하기 힘든 상황에서 프로세스 마이닝 기법을 이용하여 프로세스 모델을 도출하고 손쉽게 계획과 실적 프로세스를 비교하는 것을 목표로 한다. 따라

서 다양한 관점에서 프로세스 모델 및 속성 정보를 비교하고 각 관점별 주요 비교 인자들이 어떤 것이 있는지를 정의한다. 제안하는 비교 관점은 1) 프로세스 모델 관점의 비교, 2) 작업 관점의 비교, 3) 프로세스 인스턴스 관점의 비교, 4) 적합도 관점의 비교, 네 가지 이다. 각 관점별 주요 비교 기준 인자들은 각 절에서 자세히 설명한다. 제안하는 비교 관점별 분석 인자들을 활용하면 쉽게 계획과 실적 프로세스 및 데이터를 비교할 수 있다. 또한 이는 생산 계획 수립 시, 기준 정보로도 활용될 수 있다.

##### 4.5.1 프로세스 모델 비교

블록 조립 작업에 대한 계획 시스템의 계획 데이터와 모니터링 시스템의 실적 데이터로부터 군집별로 도출된 프로세스 모델에 대해서, 모델 차원의 직접 비교가 가능하다. 비교 인자들에 대한 정의는 <Table 3>과 같다. 두 프로세스 모델 간에 1) 일치된 노드(Matched Node), 2) 비일치된 노드(Unmatched Node), 3) 일치된 연결선(Matched Arc), 4) 비일치된 연결선(Unmatched Arc)을 비교할 수 있다. 본 사례 연구에서는 ‘작업 단계와 작업장’을 분석 액티비티로 선정하였기 때문에 노드가 일치한다는 것은 특정 작업을 해당 작업장에서 계획한 것과 실제 작업한 것이 동일하게

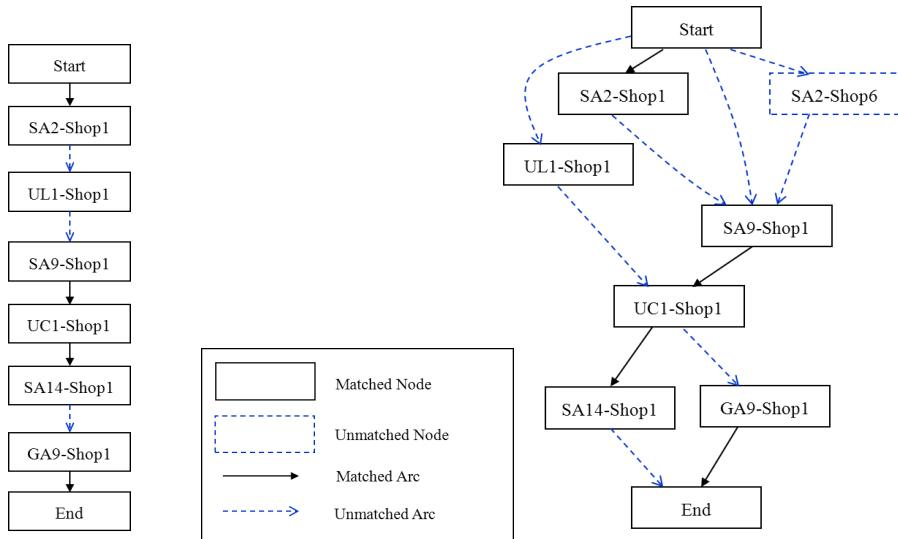
<Table 3> Comparison Factors by the Process Model Perspective

Definition of Comparison Factors	Description
1) Matched Node	Node of a processmodel exists in the other one
2) Unmatched Node	Node of a process model doesn't exist in the other one
3) Matched Arc	Arc of a process model exists in the other one
4) Unmatched Arc	Arc of a process model doesn't exist in the other one

존재한다는 것이고, 비일치 한다는 것은 대상 작업 단계가 계획에서 정의한 작업장과 다른 곳에서 작업 되었다는 것을 의미한다. 액티비티간 연결선은 작업의 흐름을 나타내는데, 일치하는 경우는 계획과 실적이 동일한 작업 순서를 가지는 것을 의미하고, 일치하지 않는 경우는 다른 작업 순서를 가지는 것을 의미한다.

프로세스 모델의 직접 비교 관점에서 'PC 5' 기준의 계획과 실적 프로세스 모델은 <Figure 6(a), (b)>와 같다. 계획 데이터 모델은 단순한

것에 비해 실적 프로세스 모델은 흐름이 다양한 것을 확인할 수 있다. 뿐만 아니라, 상호 일치된 특성과 비일치된 특성을 비교하기 용이하게 실선과 점선으로 구분했다. <Table 4>는 프로세스 모델 관점에서 비교 인자들에 대한 'PC 5'의 계획과 실적 비교 결과를 보여 준다. 작업 단계와 작업장 관점의 노드는 일치도가 높은데 비해, 작업 순서를 의미하는 연결선의 일치도는 상대적으로 낮은 것을 알 수 있다. 각 군집에 대한 결과 요약은 <Table 5>와 같다



<Figure 6(a)> Plan Model of PC 5

<Figure 6(b)> Actual Model of PC 5

<Table 4> Comparison of PC 5 by the Process Model Perspective

	Planned Model		Actual Model	
	Counts	Ratio	Counts	Ratio
1) Matched Node	8	100%	8	88.9%
2) Unmatched Node	0	0%	1	11.1%
Sum of Nodes	8	100%	9	100%
3) Matched Arc	4	57.1%	4	33.3%
4) Unmatched Arc	3	42.9%	8	66.7%
Sum of Arcs	7	100%	12	100%

<Table 5> Summary of Each Cluster by the Process Model Perspective

Cluster	Plan(P) or Actual(A) Data	Data			Process Model Perspective			
		No. of Instances	No. of Events	No. of Tasks	1) Matched Node(Ratio)	2) Unmatched Node(Ratio)	3) Matched Arc(Ratio)	4) Unmatched Arc(Ratio)
PC 1	P	48	310	13	1	0	0.667	0.333
	A	48	310	13	1	0	0.667	0.333
PC 2	P	16	68	6	1	0	0.833	0.167
	A	16	71	6	1	0	0.714	0.286
PC 3	P	15	119	30	0.467	0.533	0.282	0.718
	A	15	125	22	0.636	0.364	0.306	0.694
PC 4	P	12	108	17	1	0	0.350	0.650
	A	12	114	21	0.810	0.190	0.189	0.811
PC 5	P	8	64	8	1	0	0.571	0.429
	A	8	64	9	0.889	0.111	0.333	0.667
PC 6	P	8	66	15	0.867	0.133	0.412	0.588
	A	8	67	23	0.545	0.455	0.241	0.759
PC 7	P	8	83	16	1	0	0.227	0.773
	A	8	92	23	0.696	0.304	0.132	0.868
PC 8	P	6	34	7	1	0	0.625	0.375
	A	6	37	9	0.778	0.222	0.500	0.500
PC 9	P	4	38	14	1	0	0.571	0.429
	A	4	43	15	0.933	0.067	0.364	0.636
PC 10	P	4	35	12	0.667	0.333	0.231	0.769
	A	4	35	13	0.615	0.385	0.188	0.812
PC 11	P	4	24	6	0.667	0.333	0.200	0.800
	A	4	24	8	0.500	0.500	0.100	0.900
PC 12	P	2	8	4	1	0	1	0
	A	2	8	4	1	0	1	0

4.5.2 작업 비교

군집별 모델의 직접 비교 뿐만 아니라, 모델 내 액티비티에 해당하는 작업의 특성을

<Table 6>에 정의한 인자에 의해서 보다 자세히

비할 수 있다. 1) 동일 빈도 발생 작업 (Same Frequency Task), 2) 과실적 발생 작업

<Table 6> Comparison Factors by the Task Perspective

Definition of Comparison Factors	Description
1) Same Frequency Task	It exist as same frequency in the planned and actual process model
2) Too Much Realization Task	There are more existence in the actual model than in the planned model
3) Too Much Planned Task	There are more existence in the planned model than in the actual model
4) Unplanned Task	It doesn't exist in the planned model, but exist in the actual model
5) Skipped Task	It exists in the planned model, but doesn't exist in the actual model

(Too Much Realization Task), 3) 과계획 작업(Too Much Planned Task), 4) 계획되지 않은 작업(Unplanned Task), 5) 제외된 작업(Skipped Task) 등이 분석 인자가 될 수 있다. ‘동일 빈도 발생 작업’은 계획과 실적에서 동일한 빈도로 발생한 액티비티 즉 노드로 정의하며, ‘과실적 발생 작업’은 계획보다 실적에서 더 많이 발생한 작업으로, ‘과계획 작업’은 반대로 계획에서 더 많이 발생한 작업으로 정의한다. ‘계획되지 않은 작업’은 실적에는

〈Table 7〉 Comparison of PC 5 by the Task Perspective

	Counts	Ratio
1) Same Frequency Task	7	77.78 %
2) Too Much Realization Task	0	0 %
3) Too Much Planned Task	1	11.11 %
4) Unplanned Task	1	11.11 %
5) Skipped Task	0	0 %
Sum of Tasks	9	100 %

〈Table 8〉 Summary of Each Cluster by the Task Perspective

Cluster	Count(C) or Ratio(R)	Task Perspective				
		1) Same Frequency Task	2) Too Much Realization Task	3) Too Much Planned Task	4) Unplanned Task	5) Skipped Task
PC 1	C	13	0	0	0	0
	R	100%	0%	0%	0%	0%
PC 2	C	5	1	0	0	0
	R	83.33%	16.67%	0%	0%	0%
PC 3	C	6	8	0	8	16
	R	15.79%	21.05%	0%	21.05%	42.11%
PC 4	C	10	4	3	4	0
	R	47.62%	19.05%	14.29%	19.05%	0%
PC 5	C	7	0	1	1	0
	R	77.78%	0.00%	11.11%	11.11%	0%
PC 6	C	6	0	7	10	2
	R	24.00%	0%	28.00%	40.00%	8.00%
PC 7	C	7	2	7	7	0
	R	30.43%	8.70%	30.43%	30.43%	0%
PC 8	C	5	1	1	2	0
	R	55.56%	11.11%	11.11%	22.22%	0%
PC 9	C	11	3	0	1	0
	R	73.33%	20.00%	0%	6.67%	0%
PC 10	C	7	0	1	5	4
	R	41.18%	0%	5.88%	29.41%	25.53%
PC 11	C	4	0	0	4	2
	R	40.00%	0%	0%	40.00%	20.00%
PC 12	C	4	0	0	0	0
	R	100%	0%	0%	0%	0%



존재하는데 계획에는 존재하지 않는 작업을 의미하며, ‘제외된 작업’은 계획에는 존재하는데 실적에는 존재하지 않는 작업으로 정의한다. 본 사례 연구에서는 액티비티 즉 작업을 ‘작업 단계와 작업장’으로 정의했기 때문에 계획과 실적의 ‘작업 단계와 작업장’의 조합이 다른 경우 차이를 보이게 된다.

‘PC 5’에 대해서 작업 관점에서 분석 비교한 결과는 <Table 7>과 같다. ‘과계획 작업’과 ‘계획되지 않은 작업’이 각각 1건씩 존재하는 것을 알 수 있다. 각 군집에 대한 결과는 <Table 8>과 같이 요약 된다.

#### 4.5.3 프로세스 인스턴스 비교

군집 내에서 프로세스 인스턴스 관점에서 계획과 실적을 비교할 수 있다. 본 사례 연구의 경우 프로세스 인스턴스는 하나의 조립 블록이 만들어지기까지의 과정, 즉 프로세스가 된다. 비교 인자는 <Table 9>와 같이 1) 동일 시간 인스턴스(Same Time Instance), 2) 지연 인스턴스(Delayed Instance), 3) 공기 경과 인스턴스(Overdue Instance), 4) 지연 및 공기 경과 인스턴스(Delayed and Overdue Instance)로 정의한다. ‘동일 시간 인스턴스’는 계획과 실적의 시작과 완료 시점이 모두 동일한 프로

세스 인스턴스를 의미하며, ‘지연 인스턴스’는 처음 혹은 마지막 노드가 계획대비 지연되었지만, 작업 공기가 계획 데이터를 넘지 않은 프로세스 인스턴스로 정의 한다. ‘공기 경과 인스턴스’는 작업 공기가 계획 데이터를 넘은 경우로 정의하고, ‘지연 및 공기 경과 인스턴스’는 2)와 3)의 특성을 동시에 가지는 인스턴스로 정의한다.

<Table 10> Comparison of PC 5 by the Instance Perspective

	Counts	Ratio
1) Same Time Instance	0	0%
2) Delayed Instance	7	87.5%
3) Overdue Instance	0	0%
4) Delayed and Overdue Instance	1	12.5%
Sum of Instances	8	100%

‘PC 5’를 프로세스 인스턴스 관점에서 비교한 사례는 <Table 10>과 같다. 8개의 프로세스 인스턴스, 즉 조립 블록에 대해 7개 블록은 ‘지연 인스턴스’이고 1개는 지연과 동시에 공기도 계획 공기를 넘어선 ‘지연 및 공기 경과’ 인스턴스인 것을 확인할 수 있다. 각 군집별 결과는 <Table 11>과 같이 요약 정리 된다.

<Table 9> Comparison Factors by the Instance Perspective

Definition of Comparison Factors	Description
1) Same Time Instance	First and last node for same planned and actual instance, have the same time stamp respectively
2) Delayed Instance	First or last node was delayed, but working duration was not over plan data
3) Overdue Instance	Working duration was over plan data
4) Delayed and Overdue Instance	Both 2) and 3)

〈Table 11〉 Summary of Each Cluster by the Instance Perspective

Cluster	Count(C) or Ratio(R)	Instance Perspective			
		1) Same Time Instance	2) Delayed Instance	3) Overdue Instance	4) Delayed and Overdue Instance
PC 1	C	3	41	0	4
	R	6.25%	85.42%	0%	8.33%
PC 2	C	3	10	1	2
	R	18.75%	62.50%	6.25%	12.50%
PC 3	C	2	6	4	3
	R	13.33%	40.00%	26.67%	20.00%
PC 4	C	0	7	2	3
	R	0%	58.33%	16.67%	25.00%
PC 5	C	0	7	0	1
	R	0%	87.50%	0%	12.50%
PC 6	C	1	2	2	3
	R	12.50%	25.00%	25.00%	37.50%
PC 7	C	2	4	1	1
	R	25.00%	50.00%	12.50%	12.50%
PC 8	C	0	6	0	0
	R	0%	100%	0%	0%
PC 9	C	0	4	0	0
	R	0%	100%	0%	0%
PC 10	C	0	3	0	1
	R	0%	75.00%	0%	25.00%
PC 11	C	0	4	0	0
	R	0%	100%	0%	0%
PC 12	C	0	2	0	0
	R	0%	100%	0%	0%

4.5.4 적합도 비교

시스템의 데이터로부터 도출된 프로세스 모델은 모델의 순응도(conformance)를 평가하기 위해서 적합도(fitness)를 계산한다. 적합도는 프로세스 모델이 이벤트 로그들의 프로세스를 얼마나 잘 반영하고 있는지 수치적으로 계산한 값으로, 네트워크 N에서 자취(trace)  $\sigma$ 를 가지는 하나의 케이스에 대해서 식 (1)과 같이 적합도가 계산될 수 있다.

$$fitness(\sigma, N) = \frac{1}{2} \left( 1 - \frac{m}{c} \right) + \frac{1}{2} \left( 1 - \frac{r}{p} \right) \quad (1)$$

( $c$  : consumed tokens,  $m$  : missing tokens,  $p$  : produced tokens,  $r$  : remaining tokens)

식 (1)은 네트워크 내에서 잃어버린 토큰이 없고, 남아 있는 토큰이 없는 경우 적합도가 '1'이 됨을 알 수 있다. 반대로 소비된 토큰 모두를 잃어 버리거나, 생산된 토큰이 모두 네트워크 내에 남아 있으면 네트워크가 프로세스를 전혀 반영하고 있지 못하므로 적합도가 '0'이 된다. 즉, 식 (1)에서 첫 번째 항의 '(1-m/c)'은 성공적 실행의 비율을 의미하고, 두 번째 항의 '(1-r/p)'은 적절한 완료의 비율이 된다. 두 항의 가중치를 동일하게 1/2로 준 것이다. 이렇게 토큰의 흐름을 통해 적합도를 보다 정확하게 분석하기 위해서, 휴리스틱 마이닝을 통해 도출된 프로세스를 페트리넷 모형으로 변

<Table 12> Comparison of PC 5 by the Fitness Perspective

		c	m	p	r	Output
Plan Model	Fitness	72	0	72	0	1
	Cross fitness	58	10	58	10	0.828
Actual Model	Fitness	92	4	90	2	0.967
	Cross fitness	88	0	88	0	1

환 후 적합도를 계산 하였다.

본 연구에서 적합도 계산의 특징은 계획과 실적 프로세스의 비교를 목적으로 하기 때문에 자신의 데이터에 대한 모델의 적합도 뿐만 아니라, 상대 데이터에 대한 모델의 적합도까지 계산하고 이를 교차 적합도(cross fitness)로 정의한 것에 있다. 즉 계획 데이터 기반으

로 도출한 프로세스 모델에 실적 데이터의 적합도를 계산하는 경우와 실적 데이터 기반으로 도출한 프로세스 모델에 계획 데이터의 적합도를 계산한 경우가 각각 교차 적합도에 해당된다. 계획 데이터 기반으로 만든 프로세스 모델의 교차 적합도가 낮다는 것은 계획 데이터 관점의 프로세스 모델이 실적 데이터의 작

<Table 13> Summary of Each Cluster by the Fitness Perspective

No	Cluster	No. of Instances	Differences (Actual - Planned)		Planned Process Model		Actual Process Model	
			No. of Event	No. of Task	Fitness	Cross fitness	Fitness	Cross fitness
1	PC 1	48	0	0	1	0.959	0.939	0.939
2	PC 2	16	3	0	1	0.943	0.938	0.935
3	PC 3	15	6	-8	0.911	0.787	0.890	0.770
4	PC 4	12	6	4	1	0.750	0.932	0.805
5	PC 5	8	0	1	1	0.828	0.967	1
6	PC 6	8	1	8	0.950	0.787	0.913	0.831
7	PC 7	8	9	7	0.925	0.576	0.830	0.665
8	PC 8	6	3	2	0.846	0.646	1	0.911
9	PC 9	4	5	1	1	0.714	0.984	0.858
10	PC 10	4	0	1	1	0.750	0.977	0.818
11	PC 11	4	0	2	1	0.538	0.945	0.675
12	PC 12	2	0	0	1	1	1	1
Average					0.969	0.773	0.943	0.850

업 프로세스를 잘 반영하지 못하고 있다는 의미이고, 반대로 교차 적합도가 높다는 것은 실제 작업을 잘 반영하고 있는 모델이라는 의미이다. 또한 실적 데이터 기반의 프로세스 모델의 교차 적합도는 실적 기반의 모델이 계획을 얼마나 잘 반영하고 있는지를 나타낸다고 할 수 있다. 'PC 5'의 경우 적합도와 교차 적합도의 계산사례는 <Table 12>와 같다. 식 (1)에 의해서 소비된 토큰 개수(c), 잃어버린 토큰 개수(m), 생산된 토큰 개수(p), 남은 토큰 개수(r) 값을 이용해서 계산된다.

<Table 13>은 본 사례 연구에서 각 군집에 대해서 계획 데이터 기반 프로세스 모델과 실적 데이터 기반 프로세스 모델에 대한 적합도와 교차 적합도를 계산한 결과를 보여주고 있다. 계획 데이터 기반의 프로세스 모델과 실적 데이터 기반의 프로세스 모델을 비교해 보면 자신의 데이터에 대한 적합도 값은 계획 프로세스 모델이 평균값 기준으로 0.026(= 0.969-0.943) 높지만 교차 적합도 값은 실적 프로세스 모델이 0.077(= 0.850-0.773) 높다는 것을 확인할 수 있다. 또한 계획과 실적간의 이벤트 수와 작업 수의 차이와 적합도간의 관계를 파악할 수 있다. 적합도가 낮은 군집의 조립 블록에 대한 작업 특성을 재정의 하거나 다른 군집에 포함시키는 등의 평가 기준으로도 활용할 수 있다.

## 5. 결 론

작업장이 넓고 다양한 작업 조직과 많은 작업자가 함께 생산을 하는 조선 산업의 경우는 복잡한 생산 작업 프로세스를 잘 이해하고 관

리하는 능력이 중요한 경쟁력이 된다. 대형 조선소들은 이미 복잡한 생산 환경에 대응하기 위해 ERP(Enterprise Resource Planning) 시스템을 도입하고, APS(Advanced Planning System) 등을 개발하여 생산 관리의 수준을 높이기 위해 노력해 왔다. 그러나 조선 산업의 특성상, 복잡한 생산 프로세스와 다양한 생산 상황의 변화에 대응하여 생산을 잘 이해하고 정확한 실적을 집계하는 부분은 지속적으로 어려움이 있었다. 따라서 최근까지도 생산의 정확한 진행 상황을 모니터링 하기 위해 IT 기술의 발달과 함께 생산 공정 모니터링 시스템 개발에 주력하였고, 지금은 안정화 단계에 접어들어, 실적 데이터 기반의 많은 데이터도 실시간으로 축적 되고 있다. 따라서 이들 정보시스템의 데이터를 기반으로 보다 빠르고 효율적으로 계획과 실적 데이터를 비교하고 생산 계획 및 관리를 위한 정보로 피드백 하여 생산 관리 능력을 높이고 생산성 향상에 기여하고자 한다.

본 논문에서는 조선 산업에서 가장 기본이 되면서 중요한 조립 블록 작업에 대해서 계획과 실적 프로세스를 비교 분석 하고자 했다. 정보시스템에 존재하는 데이터를 이용하여 쉽고 빠르게 프로세스 모델을 도출하고, 다양한 관점에서 비교 분석 할 수 있는 방안을 제안하였다. 본 연구에서 제안한 분석 방법은 1) 데이터 전처리, 2) 분석 수준 정의, 3) 작업 특성을 고려한 블록 군집화, 4) 군집별 프로세스 도출, 5) 계획과 실적 프로세스 간의 비교, 다섯 단계로 구성된다. 계획 프로세스와 실적 프로세스간의 비교는 프로세스 모델 관점, 작업 관점, 프로세스 인스턴스 관점, 적합도 관점에서 비교할 것을 제안하였고, 각 관점별

주요 분석 인자를 정의하였다. 특히 적합도 관점에서는 자신의 데이터를 이용한 적합도 계산뿐만 아니라 자신의 모델에 대해서 상대 데이터를 통해 적합도를 계산하는 교차 적합도까지 정의하고 비교 분석을 수행 하였다.

본 논문에서 제안한 방법은 형상과 구조가 모두 다른 다양한 조립 블록에 대해서 쉽고 빠르게 프로세스 모델을 정의하고 계획과 실적 관점에서 작업의 특성들을 비교할 수 있는데 의의가 있다. 사례 연구를 통해서, 전반적으로 계획 단계에서는 계획자가 정의한 군집별로 거의 동일한 작업 프로세스를 계획하지만, 생산 단계에서는 상황에 따라 작업 순서 및 작업장 활용을 어느 정도 유연하게 변경하고 있는 것이 확인 되었다. 이렇게 생산에서 작업 순서 및 작업장 변경이 가능한 작업 단계들에 대해서는 계획 시에도 유연성을 두어서 보다 효율적인 계획 수립이 가능할 것으로 판단 된다. 또한 어떤 작업 특성을 가지는 군집들이 계획 대비 실적이 차이가 많이 나는지를 확인 할 수 있으므로 이를 통해 계획의 개선 방향을 정할 수 있다. 그러나 본 연구는 계획자의 경험과 판단에 의해 유사한 작업 특성을 가지는 조립 블록들을 구분한 ‘계획 케이스’를 군집으로 정의하고 활용하였기 때문에 이의 동질성(homogeneity)이 우선 보장되어야 한다. 또한 군집별로 프로세스 모델을 도출하므로 군집의 결과가 프로세스 모델에도 영향을 미친다. 따라서 계획자가 관리해 오던 ‘계획 케이스’가 작업의 유사성을 잘 반영한 것인지 군집화 방안에 대한 향후 연구가 요구된다. 결국 본 논문에서 제안한 방법은 군집의 작업 특성에 대한 동질성이 보장될 때, 계획과 실적 데이터를 이용하여 작업 프로세스

를 효과적으로 비교분석 하는 방법이라 하겠다. 제안한 방법의 타당성은 세계 최상위 그룹에 속하는 국내 대형 조선소의 계획 데이터와 실적 데이터를 이용한 사례 연구를 통해서 설명 되었다.

---

## References

---

- [1] van der Aalst, W. M. P. and Basten, T., “Inheritance of workflows : An approach to tackling problems related to change,” *Theoretical Computer Science*, Vol. 270, No. 1, pp. 125-203, 2002.
- [2] van der Aalst, W. M. P., “Business alignment : using process mining as a tool for Delta analysis and conformance testing,” *Requirements Engineering*, Vol. 10, pp. 198-211, 2005.
- [3] Cho, K. K., Oh J. S., Ryu K. R. and Choi H. R., “An integrated process planning and scheduling system for block assembly in shipbuilding,” *Annals of the CIRP*, Vol. 47, No. 1, pp. 419-422, 1998.
- [4] Cho, K. K., Sun J. G. and Oh J. S., “An automated welding operation planning system for block assembly in shipbuilding,” *International Journal of Production Economics*, Vol. 60-61, pp. 203-209, 1999.
- [5] Goedertier, S., de Weerd, J., Martens, D., Vanthienen, J. and Baesens, B., “Process discovery in event logs : An application in the telecom industry,” *Applied Soft*

- Computing, Vol. 11, pp. 1697-1710, 2011.
- [6] Hur, W. C., Bae, H., Kim S. and Jeong, K. S., "A method for business process analysis by using decision tree," The Journal of Society for e Business Studies, Vol. 13, No. 3, pp. 51-66, 2008.
- [7] Vullers, Jansen M. H., van der Aalst, W. M. P., and Rosemann, M., "Mining configurable enterprise information systems," Data and Knowledge Engineering, Vol. 56, No. 3, pp. 195-244, 2006.
- [8] Jung, J. Y., "PROCL : A process log clustering system," The Journal of Society for e Business Studies, Vol. 13, No. 2, pp. 181-194, 2008.
- [9] Lee, D. and Bae H., "Analysis framework using process mining for block movement process in shipyards," ICIC Express Letters, Vol. 7, No. 6, pp. 1913-1917, 2013.
- [10] Lee, S., Kim B., Huh M., Cho S., Park S. and Lee D., "Mining transportation logs for understanding the after assembly block manufacturing process in the shipbuilding industry, Expert Systems with Applications," Vol. 40, No. 1, pp. 83-95, 2013.
- [11] de Medeiros, A. K. A., Weijters, A. J. M. M. and van der Aalst, W. M. P., "Genetic Process Mining : An Experimental Evaluation, Data Mining and Knowledge Discovery," Vol. 14, No. 2, pp. 245-304, 2007.
- [12] Rozinat, A. and van der Aalst, W. M. P., "Decision Mining in ProM, Proc," 4th Int. Conf. on Business Process Management, pp. 420-425, 2006.
- [13] Song, M., Günther C. W. and van der Aalst, W. M. P., "Trace clustering in process mining," BPM 2008 Workshops, Lecture Notes in Business Information Processing, Vol. 17, pp. 109-120, 2009.
- [14] de Weerd, J., Schupp, A., Vanderloock, A. and Baesens, B., "Process Mining for the multi faceted analysis of business processes - A case study in a financial services organization," Computer in Industry, Vol. 64, pp. 57-67, 2013.
- [15] Weijters, A. J. M. M., van der Aalst, W. M. P. and de Medeiros A. K. A., "Process Mining with Heuristics Miner Algorithm," BETA Working Paper Series, WP 166, Eindhoven University of Technology, Eindhoven, 2006.

## 저 자 소개



이동하  
2000년  
2002년  
현재  
관심분야

(E-mail : dongha@dsme.co.kr)  
부산대학교 산업공학과 (석사)  
부산대학교 산업공학과 (박사수료)  
대우조선해양 중앙연구소 생산시스템연구그룹 과장  
생산계획 및 관리 시스템, 프로세스 마이닝, 데이터마이닝,  
물류관리 등



박재훈  
2008년  
2013년  
현재  
관심분야

(E-mail : pjh3479@pusan.ac.kr)  
부산대학교 산업공학과 (석사)  
부산대학교 산업공학과 (박사)  
부산대학교 산업공학과박사 후 연구원  
프로세스 효율성, 벤치마킹, DEA(Data Envelopment  
Analysis), 프로세스 마이닝 등



배혜림  
1998년  
2002년  
현재  
관심분야

(E-mail : hrbae@pusan.ac.kr)  
서울대학교 산업공학과 (석사)  
서울대학교 산업공학과 (박사)  
부산대학교 산업공학과 부교수  
프로세스 마이닝, BPM(Business Process Management),  
Multi organizational Process Management in e-Business  
Environments, Enterprise Information System Integration,  
e-SCM and e-CRM 등