

# 의사결정나무를 활용한 비즈니스 프로세스 분석

## A Method for Business Process Analysis by using Decision Tree

허원창(Hur, Wonchang)\*, 배혜림(Bae, Hyerim)\*\*, 김 승(Kim, Seung)\*\*\*,  
정기성(Jeong, Kiseong)\*\*\*\*

### 초 록

기업이 비즈니스 프로세스의 중요성을 인식하게 되면서 비즈니스 프로세스 관리시스템에 대한 관심이 높아졌다. 그러나 전통적으로 BPMS는 업무프로세스의 모형화 및 실행에 그 초점이 맞추어져 있어, 상대적으로 프로세스 관리의 궁극적인 목적인 프로세스 효율성 제고나 더 나아가 프로세스를 혁신하는 데에는 큰 기여를 하지 못하였다. BPMS는 일반적으로 방대한 량의 이력 데이터를 남기게 되는데, 이에는 프로세스의 실행과 관련된 다양한 의미 있는 정보나 규칙들이 숨어있다. 본 연구는 데이터마이닝 기법을 활용하여 프로세스 이력 데이터를 분석하는 체계적인 방법론을 제시한다. 제시된 방법론은 프로세스 분석을 위한 데이터 모형, 분석절차, 그리고 시스템 프레임워크를 포함한다. 또한 그 유용성을 검증하기 위하여, 시뮬레이션 데이터를 이용한 실험결과를 제시한다.

### ABSTRACT

The Business Process Management System (BPMS) has received more attentions as companies increasingly realize the importance of business processes. However, traditional BPMS has focused mainly on correct modeling and exact automation of process flow, and paid little attention to the achievement of final goals of improving process efficiency and innovating processes. BPMS usually generates enormous amounts of log data during and after execution of processes, where numerous meaningful rules and patterns are hidden. In the present study we employ the data mining technique to find out useful knowledge from the complicated process log data. A data model and a system framework for process mining are provided to help understand the existing BPMS. Experiments with the simulated data demonstrate the effectiveness of the model and the framework.

**키워드** : 비즈니스 프로세스 관리, 비즈니스 프로세스 분석, 데이터마이닝  
Business Process Management, Business Process Analysis, Data Mining

---

이 논문은 2006년~2007년도 인하대학교의 지원에 의하여 연구되었음(INHA-35028-1).

\* 인하대학교 경영학부 조교수

\*\* 교신저자, 부산대학교 산업공학과 조교수

\*\*\* 서울대학교 산업공학과

\*\*\*\* ㈜네오플렉스 부장

2008년 02월 21일 접수, 2008년 05월 20일 심사완료 후 2008년 07월 23일 게재확정.

## 1. 서 론

기업환경의 급속한 변화에 따라 오늘날 기업들은 안팎으로부터의 많은 도전에 직면해 있다. 제품과 서비스의 종류가 다양해 지면서 기업내의 업무 프로세스는 그 수가 증가하고 복잡해졌으며, 짧아진 수명주기는 프로세스의 잦은 변경을 요구하고 있다. 이러한 변화속에서, 1990년대 초반부터 등장했던 비즈니스 프로세스 혁신(BPR; Business Process Reengineering), 혹은 지속적 프로세스 개선(CPI; Continuous Process Innovation) 등의 경영혁신 개념이 다시 새롭게 주목 받게 되었고 비즈니스 프로세스 관리(BPM; Business Process Management)에 관한 연구가 활발히 이어지고 있다[3, 6].

비즈니스 프로세스 관리를 지원하는 많은 상용 BPM 시스템들이 등장하였다. 이들은 업무 프로세스의 자동화된 수행과 그 과정의 통제 및 관리에 중점을 두고 설계되어 왔다 [9, 10]. 그러나 비즈니스 프로세스 관리의 궁극적인 목적은 비즈니스 프로세스의 감시와 통제를 통해 업무 효율을 증진하고 나아가 프로세스를 혁신하는 것이라 할 수 있다 [4, 7]. 효과적인 프로세스 혁신은 체계적인 분석에 기반한다. 따라서 업무 프로세스 혁신 과정에는 어떠한 데이터를 어떠한 도구와 절차로 분석해야 하는가에 대한 체계적인 접근방법이 요구되고 있다.

본 연구에서는 많은 양의 데이터를 분석할 수 있고, 또한 비교적 간단한 방법으로 유의미한 정보를 찾을 수 있는 의사결정나무 기법을 이용하여 비즈니스 프로세스를 효과적으로 분석하고 평가할 수 있는 체계적인 방

법론을 제시하고자 한다. 본 연구는 특히 비즈니스 프로세스의 수행도 지표의 단순한 측정에 그치지 않고, 다양한 프로세스의 속성 및 관련 수행도 지표들 간의 인과관계를 체계적으로 추적할 수 있는 방법론의 설계에 중점을 두었다. 또한 제시된 분석 방법의 타당성 입증에 위하여 가상의 비즈니스 프로세스를 모형화하고, 시뮬레이션을 통하여 프로세스 데이터를 생성하여 분석에 적용하였다.

## 2. 관련 연구

비즈니스 프로세스의 분석 및 평가와 관련하여 비교적 많은 연구들이 BAM(Business Activity Monitoring)과 관련한 주제를 다루고 있다[4, 7]. 특히 기업의 성과지표(KPI, Key Performance Indicator)와 관련하여 비즈니스 프로세스 분석을 활용한 연구들이 돋보이는데[12, 13], Srinivasan et al.은 IBM의 공급사슬 관련 업무에서 비즈니스 프로세스의 수행도를 평가하는 지표들을 정의하고 웹 로그 데이터를 분석하여 평가하는 방법을 제시하였다. 하지만, 제시된 지표는 문서의 생성시간, 손실문서의 개수 등 프로세스의 수행도를 파악하기 위한 비교적 쉽게 계산될 수 있는 것들을 사용하였다[12]. Hwang et al.은 균형성과지표와 BPR을 연계시킬 수 있는 방법론을 제안하였고[8], Kung et al.은 대형은행의 성과지표 측정을 위한 BPM의 활용에 대한 사례연구를 수행하였다[11].

비즈니스 프로세스의 개선을 위한 엄밀한 분석기법에 대한 연구는 시뮬레이션이나 시계열 분석을 이용한 분석에 관한 연구들을 일

부 찾을 수 있을 뿐, 그 중요성에도 불구하고 비즈니스 프로세스의 실행 및 통제나 BAM에 관한 연구에 비해 상대적으로 부족하다고 할 수 있다[1]. 특히 프로세스 모형이 가진 구조적 문제를 진단하고 이를 효과적으로 개선하기 위한 목적으로 설계된 분석방법론에 관한 연구는 찾아보기 힘들다. 이는 대부분의 BPR 기법이 시스템화된 절차에 의하기 보다는 전문가의 주관적인 지식과 경험에 의존하는 경향이 많기 때문이다. 다만 이와 관련하여 가장 주목할만한 연구는 비즈니스 프로세스 인텔리전스(BPI, Business Process Intelligence)에 관한 연구이다. 이 연구에서 저자들은 비즈니스 인텔리전스(BI, Business Intelligence)의 개념을 비즈니스 프로세스 분석에 활용할 수 있는 가능성을 제시하였으나, 주로 개념적인 정립을 중심으로 단편적인 분석사례만을 제시하고 있어 체계화된 분석절차 및 사례가 부족하다고 할 수 있다[5].

### 3. 비즈니스 프로세스 분석 방법론

#### 3.1 데이터 개체와 수행도 지표

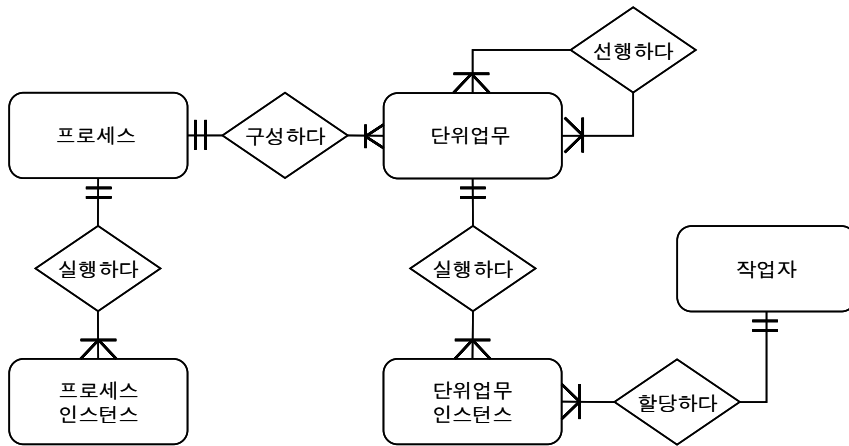
비즈니스 프로세스의 분석은 비즈니스 프로세스의 실행 이력이 기록된 데이터를 대상으로 수행된다. 비즈니스 프로세스의 실행 이력이 저장되는 데이터 스키마는 BPM 시스템에 따라 차이가 있으나 일반적으로 프로세스, 단위업무, 그리고 작업자를 주요한 개체로 하여 모델링 된다. <그림 1>은 본 연구에서 참조한 상용 BPM 시스템 및 WFMC의 참조모형을 바탕으로 분석에 사용될 비즈

니스 프로세스 실행이력 데이터의 데이터 스키마를 개략적으로 모델링 한 것이다. 본 연구의 분석방법은 이와 같은 데이터 스키마를 바탕으로 설계 되었다.

그림에서 보는 바와 같이 비즈니스 프로세스는 프로세스 개체와 이를 구성하는 단위업무 개체들로 표현된다. 단위업무 개체들 간에는 업무의 수행순서를 나타내는 선후관계가 정의된다. 프로세스 인스턴스 개체는 하나의 프로세스 정의가 실행될 때마다 새롭게 생성되는 개체로서 프로세스의 실행 과정에서 발생하는 각종 이벤트의 발생 시점 및 그에 따른 프로세스의 상태 변화 등을 기록, 관리하기 위해 사용되는 데이터 개체이다. 이와 마찬가지로 프로세스를 구성하는 개별 단위업무에 대해서도 프로세스에 정의된 실행순서에 따라 단위업무 인스턴스 개체를 생성하여 그 실행 이력을 기록, 관리한다. 단위업무 인스턴스 개체가 생성되면 BPM 시스템은 정의된 비즈니스 규칙에 따라 해당 업무를 담당할 작업자 개체를 할당한다.

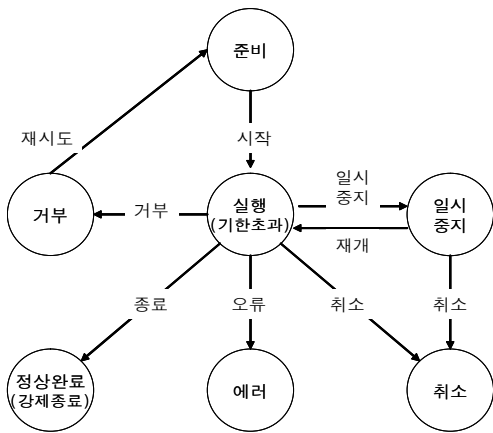
프로세스 인스턴스와 단위업무 인스턴스는 구체적으로는 <그림 2>와 같은 상태전이 관계에 따른 각 개체의 상태변화 이력을 기록하게 된다. 이러한 상태전이 관계는 BPM 시스템이 비즈니스 프로세스의 실행 과정을 통제하기 위하여 발생 시키는 ‘이벤트’와 이에 대응되는 프로세스 및 단위업무 데이터 개체들이 어떠한 상태로 전이하게 되는지를 도식화 한 것이다.<sup>1)</sup> 일반적으로 개별 단위업무들은 최초 ‘준비’ 상태에 머무르고 있다

1) 이러한 데이터 개체의 상태전이 관계는 BPM 시스템이 어떠한 메커니즘에 따라 프로세스의 실행을 통제하느냐에 따라 차이가 있을 수 있다.



〈그림 1〉 비즈니스 프로세스 데이터베이스 스키마

가 업무가 실행되면 ‘실행’ 상태로 전이하게 된다. 이어 업무가 정상적으로 완료되면 ‘정상완료’ 상태로 전이하게 되거나, 업무의 진행 상황에 따라 ‘일시중지’, ‘취소’, ‘에러’, ‘거부’ 등의 예외적인 상태로 전이하기도 한다.



〈그림 2〉 구성업무의 상태변이

이와 같은 상태전이 관계의 변화 이력은 비즈니스 프로세스와 관련된 개별 데이터 개체들의 수행도를 측정하는 데 주요하게 사용

된다. <표 1>은 각 데이터 개체의 수행도 측정에 사용될 수 있는 여러 가지 수행도 지표들을 예시하고 있다. 표에는 분석방법의 효과적인 설명을 위하여 각 개체와 수행도 지표들을 대상으로 별도의 기호를 정의하였다. 가령 제 4.1절에서 사례로 제시된 ‘보험가입처리’ 프로세스를  $P_1$ 로 표현하고 그 수행도 지표 중 하나인 ‘수행시간’을  $PP_1$ 으로 표현할 경우 ‘ $PP_1(P_1)$ ’은 보험가입처리 프로세스의 수행시간을 나타낸다.

한편, 각 데이터 개체들은 여러 가지 다양한 속성을 가지게 된다. <표 1>은 각 개체가 가질 수 있는 속성들을 함께 예시하고 있다. 각 개체 별로 정의될 수 있는 속성은 개체의 종류에 관계없이 공통적으로 해당되는 속성과 특정한 개체 인스턴스에 고유하게 정의될 수 있는 속성으로 구분될 수 있다. 가령, 모든 프로세스는 종류에 관계없이 관리자, 긴급여부, 종료시한 등의 속성을 가질 수 있지만, ‘보험가입처리’ 프로세스는 ‘보험접수 타입’, ‘보험가입자 이름’ 등 해당 프로세스에

〈표 1〉 데이터 개체 별 속성과 수행도 지표 예시

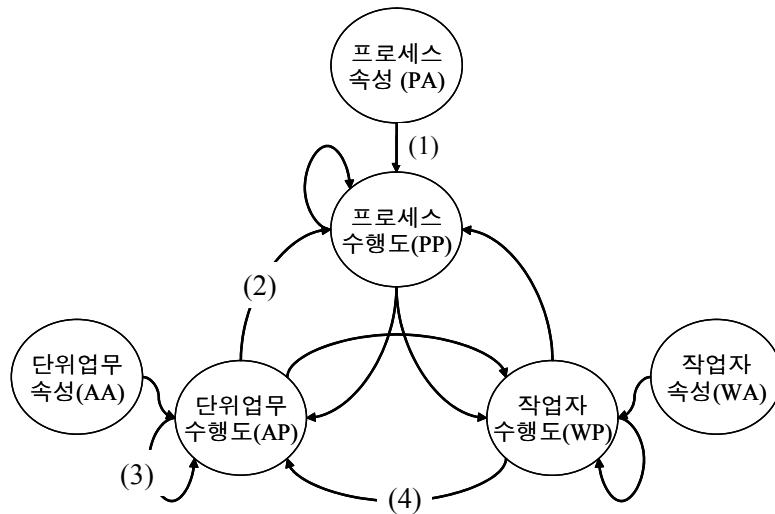
개체	속성과 수행도 지표	
프로세스( $P_i$ )	프로세스 속성 (Process Attribute) : $PA(P_i)$	<ul style="list-style-type: none"> <li>프로세스 관리자/생성자, 상위/하위 프로세스, 긴급여부, 우선순위, 종료시한, 프로세스 생성자, 시작시간, 완료시간, 소요기간, 정지기간, 완료상태, 첨부문서 개수, 반려회수 등</li> <li>프로세스 실행 시에 정의되는 해당 프로세스 고유 데이터들 Ex) 보험접수 타입, 가입프로세스 대상 고객의 신상정보 등</li> </ul>
	프로세스 수행도 (Process Performance) : $PP_i(P_i)$	<ul style="list-style-type: none"> <li>수행시간, 기한대비 수행시간, 기한 내 완료여부, 완료상태, 이상상태 발생여부 등</li> </ul>
단위업무( $A_i$ )	구성업무 속성 (Activity Attributes) : $AA(A_i)$	<ul style="list-style-type: none"> <li>선/후행업무 수, 분기/병합 타입, 담당부서, 담당자, 후행조건, 시작시간, 소요기간, 정지기간, 완료시간, 수행결과, 긴급여부, 우선순위, 종료시한 등</li> <li>업무의 실행 시에 정의되는 해당 단위업무 고유 데이터들 Ex) 신용조회를 위한 고객정보, 계좌조회를 위한 계좌정보 등</li> </ul>
	단위업무 수행도 (Activities Performance) : $AP_i(A_i)$	<ul style="list-style-type: none"> <li>수행시간, 기한대비 수행시간, 기한 내 완료 비율, 성공완료 비율, 이상상태 발생 비율, 반복수행 횟수 등</li> </ul>
작업자( $W_i$ )	작업자 속성 (Worker Attributes) : $WA(W_i)$	<ul style="list-style-type: none"> <li>직위, 나이, 역할, 직책, 관리자, 부서, 권한, 직능, 주요담당업무, 주요 담당 프로세스 등</li> </ul>
	작업자 수행도 (Worker Performance) : $WP_i(W_i)$	<ul style="list-style-type: none"> <li>예) 업무처리 시간, 정상처리 여부, 기한 내 처리여부 등</li> </ul>

서만 정의되는 고유한 속성들을 가질 수 있다. 이러한 속성들은 WFMC 참조모형에서는 워크플로우 관련데이터(Workflow relevant data)로 지칭하고 있으며[14], 일부 상용 BPM 시스템에서는 ‘프로세스 변수’와 같은 명칭을 사용하여 프로세스 모델링 시 별도로 정의할 수 있도록 하고 있다.

### 3.2 수행도 지표의 분석절차

지금까지 비즈니스 프로세스 분석에 사용될 주요 데이터의 구조와 주요한 데이터 개체 및 그 속성, 그리고 분석의 직접적인 대상

이 되는 각 개체의 수행도 지표들에 대하여 설명하였다. 본 연구에서 제시하고자 하는 분석방법론의 목적은 이와 같은 수행도 지표들의 일차적인 측정에 머무르지 않고, 그 지표가 어떠한 원인에 의하여 영향을 받게 되었는가를 파악할 수 있도록 하는 것이다. 가령, 수행도가 낮게 측정되는 프로세스가 있다면 그 원인이 특정한 단위업무가 지체됨에서 비롯된 것인지, 혹은 프로세스와 관련된 특정한 속성의 차이에 의한 영향 때문인지, 아니면 특정한 작업자나 관리자의 낮은 수행도 때문인지를 체계적으로 추적할 수 있도록 하는 것이다.



〈그림 3〉 프로세스, 단위업무, 작업자 수행도의 인과관계 다이어그램

이러한 분석방법의 설계를 위해서는 먼저 분석의 대상이 되는 개체들의 수행도 지표들이 어떠한 인과관계로 연관되어 있는지에 대한 정의가 필요하다. 앞서 살펴본 비즈니스 프로세스의 데이터 구조를 바탕으로 본 연구에서는 <그림 3>과 같은 수행도 지표들간의 인과관계를 정의하였다. 그림에서 보는 바와 같이 본 연구의 주요 분석 대상인 세 가지 데이터 개체의 수행도 지표는 우선 개별 속성의 차이에 의해 영향을 받게 된다. 또한 각 수행도 지표들 간에는 상호간의 인과관계가 존재하는 것으로 표현하였다. 특히, 동종의 데이터 개체들 사이에도 수행도의 인과관계가 존재하는 것으로 표현하였는데, 가령 같은 프로세스 개체라 하더라도 서로 다른 비즈니스 프로세스는 서로의 수행도 지표간에 인과관계가 존재할 수 있음을 표현하였다. 따라서 그래프에서 각 개체의 속성(PA, AA, WA)은 인과관계와 관계없이 결정되는 외생

(exogenous)변수라 할 수 있으며, 그 수행도(PP, AP, WP)는 다른 요인에 따라 변화하게 되는 내생(endogenous)변수라 할 수 있을 것이다. 예를 들어, <그림 1>은 프로세스의 수행도(PP)가 프로세스가 가지는 특정한 속성(PA)의 영향에 의해 변화될 수 있다는 것을 의미한다.<sup>2)</sup>

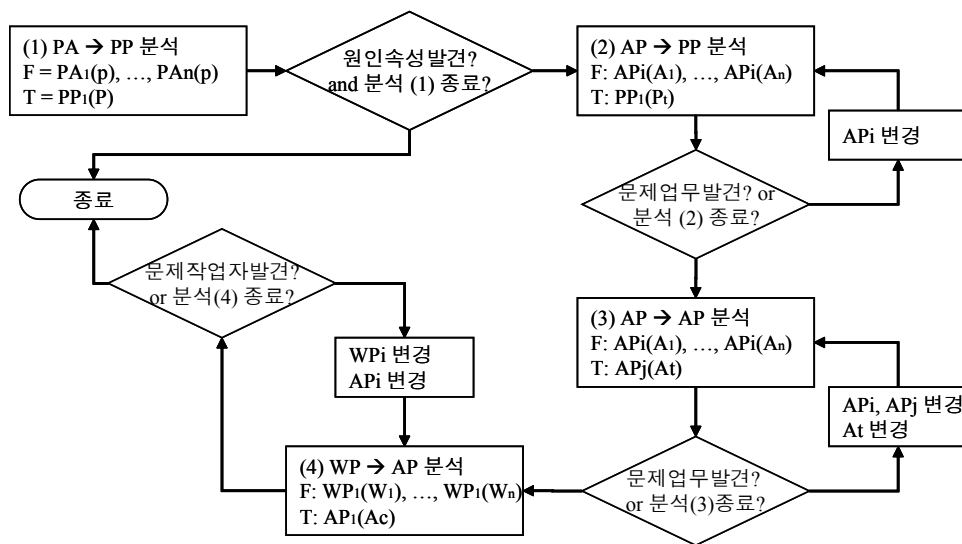
비즈니스 프로세스의 분석은 관리자가 관심을 가지는 주요한 수행도 지표를 계측하고 모니터링 하는 것으로부터 시작된다. 수행도 지표의 계측 및 모니터링은 대부분의 BPM 시스템에 의하여 제공되는 기능이다. 이를 바탕으로 사용자는 대상 수행도 지표에 영향을 주는 다른 요인들을 발견하기 위하여 <그림 3>의 다이어그램을 바탕으로 인과관계 탐색경로를 설정한다. 가령 ‘보험가입처

2) 기본적으로 모든 내생변수는 서로간의 인과관계를 가질 수 있다고 가정하였다. 다만, 분석할 프로세스의 종류에 따라 일부 인과관계는 의미를 가지지 않을 수도 있다.

리' 프로세스의 수행도 지표인 프로세스 수행시간이 늦어지는 문제를 분석하고자 한다면, <그림 1~4>의 경로를 주요한 탐색경로로 설정하고 이 경로에 따라 프로세스의 수행도 지표에 영향을 주는 속성의 파악, 프로세스 수행도에 영향을 주는 단위업무의 수행도의 파악, 해당 단위업무의 수행도에 영향을 주는 다른 단위업무의 수행도, 그리고 이에 영향을 주는 작업자의 수행도 지표 등의 순서로 인과관계를 탐색해 나갈 수 있다.

<그림 4>는 이와 같은 탐색과정을 보다 구체적으로 제시하였다. 사용자는 (1)의 인과관계의 분석단계에서 총 n개의 프로세스 속성  $PA_1, PA_2, \dots, PA_n$ 을 입력변수(F)로 설정하고,  $PP_1$ 을 종속변수(T)로 설정한 후 인과관계 분석을 수행한다. 그 결과로  $PP_1$ 의 차이에 영향을 주는 여러 가지 프로세스 속성들 중 그 영향력이 큰 순서대로 주요 요인들을 파악할 수 있다. (1)의 분석에서 주요한

요인이 발견되지 않는다면, 이후 사용자는 (2)의 분석단계에서  $P_1$ 을 구성하는 n개의 단위업무의 수행도  $AP_1(A_1), \dots, AP_1(A_n)$ 를 입력변수로 설정하고,  $PP_1$ 을 종속변수로 설정하여 인과관계 분석을 수행할 수 있다. 이를 통해 사용자는 어떤 단위업무의 수행도 지표가  $PP_1(P_1)$ 의 차이에 가장 큰 영향을 주는지를 파악할 수 있다. 만약 다른 수행도 지표를 고려하고 싶다면 수행도 지표를 변경하여 해당 단계의 분석을 반복할 수 있다. 같은 방식으로 (3) 단계에서는 (2) 단계에서 파악된 특정한 단위업무를 대상으로 해당 단위업무의 수행도 지표에 영향을 줄 수 있는 다른 단위업무들의 수행도 지표를 입력변수로 설정하여 분석한다. 마지막으로 (4) 단계 분석을 통해 이전 단계에서 주요한 요인으로 파악된 단위업무의 수행도 지표들이 작업자들의 수행도 지표에 의해 어떠한 영향을 받게 되는지를 파악할 수 있다.



<그림 4> 분석절차 예시

개별 인과관계의 분석을 위해서 의사결정 나무(Decision Tree) 분석을 사용하였다. 의사결정나무 분석은 다수의 입력변수 중 종속 변수의 차이를 가장 잘 설명하는 입력변수를 기준으로 데이터 집합을 반복적으로 분할함으로써 입력변수의 종속변수에 대한 영향력을 파악할 수 있는 통계기법이다. 또한, 의사결정나무는 비모수 통계기법이기 때문에 각 변수들의 분포에 대한 추정이 필요하지 않으며, 여러 개의 요인을 동시에 고려하여 가장 영향도가 높은 요인을 순차적으로 파악할 수 있기 때문에 본 연구가 제시하는 분석방법론과 같이 다수의 입력변수를 사용한 탐색적 분석이 필요한 경우에 적합한 통계기법이라고 할 수 있다[2].

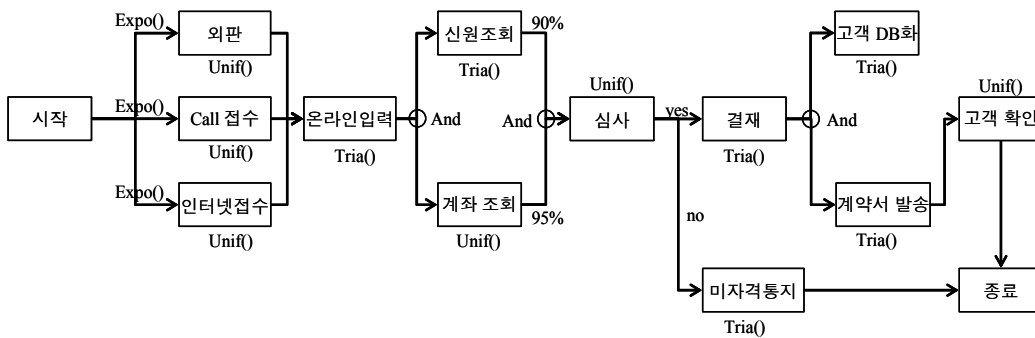
#### 4. 실험

##### 4.1 시뮬레이션 데이터

본 절에서는 앞서 설명한 분석절차에 따라 실제로 비즈니스 프로세스 이력데이터의 분석을 수행하였다. 이를 위하여 <그림 5>와

같이 가상의 보험가입처리 프로세스를 표현하는 시뮬레이션 모형을 설계하고 이를 이용하여 분석에 사용될 이력데이터를 생성하였다. 각 단위업무에 표시된 기호는 해당 업무의 처리시간의 분포를 나타내는데, Tria()는 삼각분포, Unif()는 균일분포, 그리고 Expo()는 지수분포를 나타낸다.

설계된 시뮬레이션 모형은 보험회사에서 신규가입을 처리하는 비즈니스 프로세스를 묘사한다. 프로세스는 고객이 외판원, 전화, 또는 인터넷을 통하여 가입신청을 하면서부터 시작된다. 담당자가 신청내용을 온라인 상에 입력하면 신청자의 개인 정보가 입력된 문서가 조회 부서로 넘어가게 된다. 이후 신청자의 신원조회와 금융계좌 정보의 조회가 동시에 진행된다. 두 조회 업무의 결과가 통보되어야만 이후 심사업무를 시작할 수 있다. 심사업무를 통해 담당자는 이전의 조회 결과를 검토하고 그에 따라 가입처리 여부를 심사한다. 심사에 통과되지 못한 경우에는 미 자격 통지 업무를 거친 후 프로세스가 종료하게 된다. 심사를 통과하게 되면 결제업무를 거쳐 신규고객을 고객 DB에 입력하는 업무와 계약서를 발송하는 업무가 동시에 진



<그림 5> 보험가입 프로세스



행된다. 고객이 발송된 계약서를 확인하게 되면 본 프로세스는 종료하게 된다.

분석의 정확성을 검증하기 위하여 시뮬레이션 모형의 인수들을 조정하여 분석을 통해 밝혀내야 할 문제 상황들을 다음과 같이 설정하였다.

- 1) 우선, 하루 중 처리되는 보험가입 업무의 접수가 오전에 몰리는 상황을 모델링하기 위하여 보험가입처리 신청이 오전(9:00~12:00), 오후(13:00~18:00)에 각각 50%씩 접수되도록 하였다.
- 2) 프로세스의 구성 업무 중 AND 분기와 병합 흐름의 사이에 수행되는 '신원조회' 업무와 '계좌조회' 업무가 전체 프로세스 흐름에 있어서 약간의 병목이 걸리도록 분포를 조절하였다.
- 3) '신원조회' 업무를 처리하는 업무담당자

를 작업자 1, 작업자 2, 작업자 3의 3인으로 설정하고 작업자 1의 업무처리 속도가 가장 빠른 것으로 설정하였다.

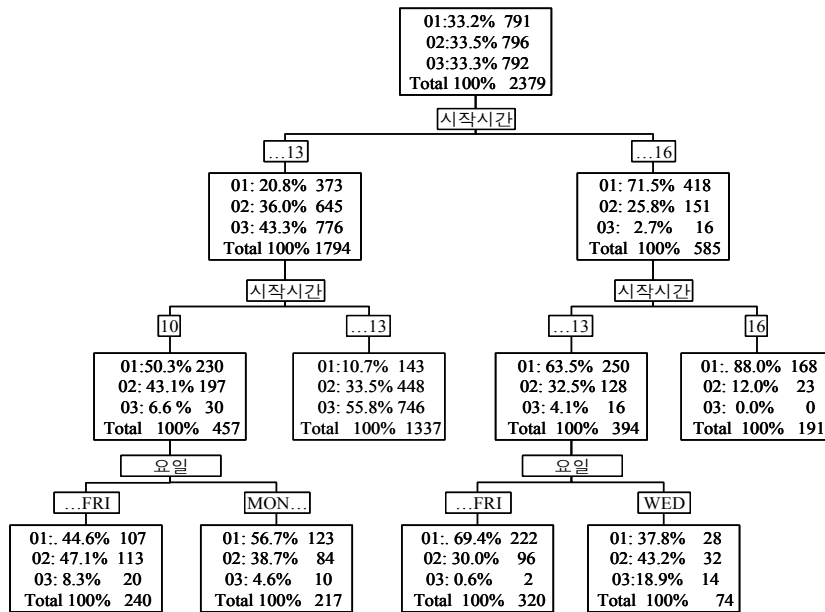
- 4) '신원조회' 업무에서 발생하는 병목으로 인해 '결재' 업무의 수행시간이 영향을 받는 상황을 모형화 하기 위하여, '심사' 업무의 수행시간의 분포를 '신원조회' 업무의 수행시간에 일부 종속되도록 설정하였다.

### 4.2 분석결과

이상의 데이터에 대하여 제 3.2절에 제시된 분석절차에 따라 분석을 수행하였다.

#### 4.2.1 프로세스 속성과 수행도 지표의 연관관계 분석(PA → PP)

첫 번째 실험단계에서는 '보험가입처리' 프



〈그림 6〉 프로세스 속성에 따른 프로세스 수행도 변화

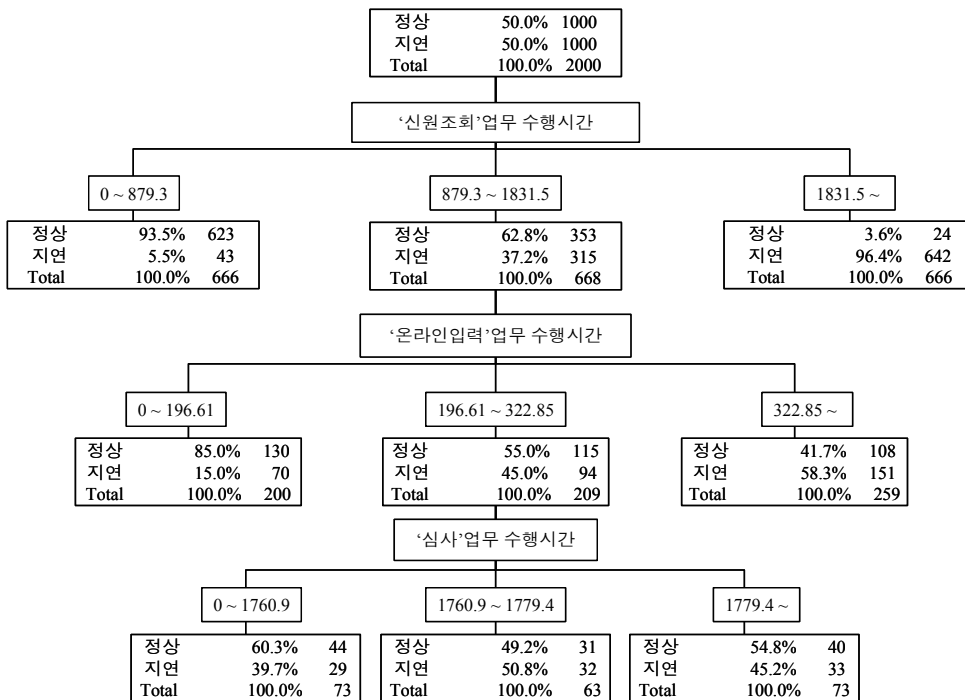
로세스의 수행도에 영향을 주는 프로세스 속성을 파악하였다. 분석의 입력변수로는 프로세스의 기초적인 속성 이외에, 프로세스 시작시간, 프로세스 시작요일 등의 속성변수를 추가하여 분석을 수행하였다. 분석결과 <그림 6>과 같은 의사결정나무를 얻었다. 총 2379개의 프로세스 인스턴스를 대상으로 분석이 수행되었으며, 프로세스의 수행시간은 크게 3개의 계열(01 : 빠름, 02 : 평균, 03 : 느림)로 분류하였다.

그림에서 보는 바와 같이 ‘보험가입처리’ 프로세스의 수행시간은 1차적으로 프로세스의 시작시간에 따라 차이가 나는 것으로 나타났다. 즉 13시를 기준으로 그 이전에 시작된 프로세스와 그 이후에 시작된 프로세스들 간의 수행시간의 차이가 극명하게 나타났다.

예를 들어 16시 이후에 시작된 보험가입 처리 프로세스 중 88%가 평균보다 빨리 끝났으나 (노드 X), 10시~13시에 시작된 프로세스의 89.3 %는 평균 혹은 그 보다 더 오랜 시간이 소요되는 것으로 파악되었다. 이러한 결과는 본 실험을 위해 생성한 시뮬레이션 데이터의 생성조건과도 일치하는 것임을 알 수 있다.

#### 4.2.2 단위업무 수행도와 프로세스 수행도의 연관관계 분석(AP → PP)

두 번째 단계에서는 단위 업무의 수행시간과 프로세스의 수행시간 사이의 연관관계를 파악하기 위한 분석이 진행되었다. 분석을 위하여 ‘보험가입처리’ 프로세스를 구성하는 모든 단위업무의 수행시간을 입력변수로 설정하여 의사결정나무 분석을 시행하였다. 그 결



<그림 7> 단위업무 수행시간에 따른 프로세스 수행시간의 변화

과가 <그림 7>에 제시되었다.

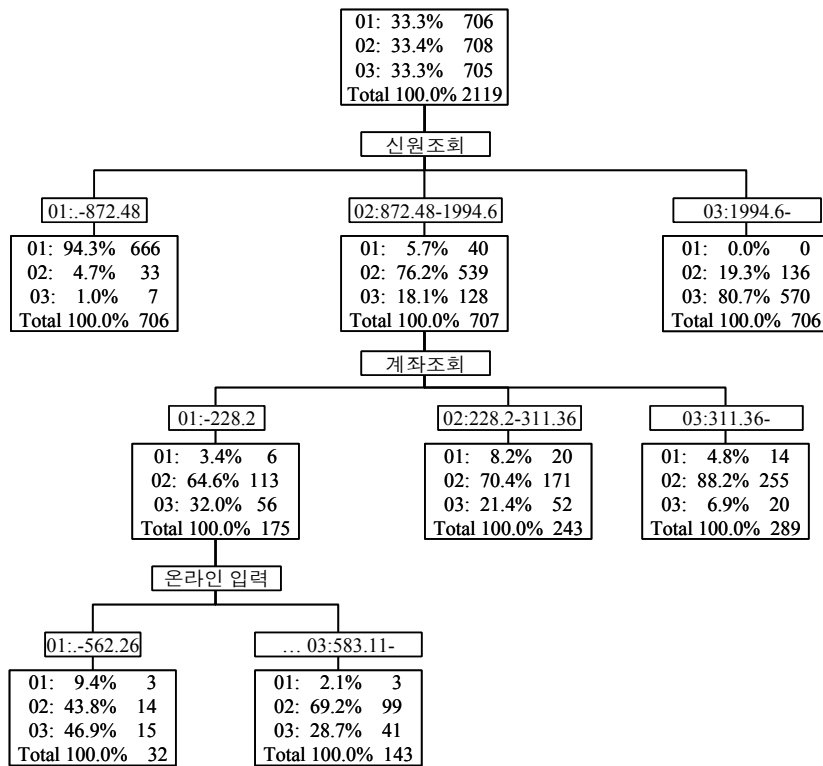
그림에서 보는 바와 같이 프로세스의 수행 시간은 ‘신원조회’ 업무의 수행시간에 따라 크게 차이가 나는 것을 알 수 있다. ‘신원조회’ 업무가 빨리 완료되었을 경우에는 그 중 93.5%의 프로세스가 정상시간 내에 종료하였으나(노드 X), ‘신원조회’ 업무가 지연되었을 경우에는 반대로 96.4%의 경우에서 프로세스 역시 지연되는 것으로 나타났다(노드 Y). 이러한 결과는 시뮬레이션 모형에서 설정한 사항(가정 2)와 일치하는 결과이다.

‘신원조회’ 업무에 이어서 두 번째로 프로세스의 수행시간에 영향을 주는 단위업무로 ‘온라인입력’ 업무가 파악되었다. ‘온라인입력’

업무의 수행시간은 시뮬레이션 모형에서 그 수행시간을 독립적으로 설정하였으나, 프로세스의 구조 상 병목의 지점에 위치함으로 인해 이러한 결과가 도출된 것으로 판단된다.

#### 4.2.3 단위업무들 간의 수행도 연관관계 분석(AP → AP)

다음 단계에서는 (2)에서 파악된, ‘신원조회’, ‘온라인입력’, 그리고 ‘심사’ 세 가지 업무를 차례로 종속변수로 설정하고, 나머지 단위업무들의 수행시간을 입력변수로 설정하여 단위업무의 수행시간 사이의 연관관계를 파악하고자 하였다. <그림 8>은 이 단계에서 얻어진 의사결정나무 중 ‘심사’ 업무를 대상



<그림 8> 선행업무들의 수행도에 따른 ‘심사’ 업무의 수행시간 변화

으로 수행된 분석결과를 보여준다.

그림에서 보는 바와 같이 ‘심사’ 업무의 수행시간은 ‘신원조회’ 업무의 수행시간에 따라 큰 영향을 받는 것으로 나타났다. 노드 X와 Y에서 보는 바와 같이 ‘신원조회’ 업무가 평균보다 늦게 완료된 총 706개의 프로세스 인스턴스 중에서 80.7%인 570개의 프로세스 인스턴스에서 ‘심사’업무가 늦게 완료되었음을 알 수 있다(계열 03). 반대로 ‘신원조회’ 업무가 빨리 종료된 경우의 94.3%에서 ‘심사’ 업무 역시 빨리 완료되었음을 알 수 있다.

#### 4.2.4 작업자 수행도와 단위업무 수행도의 연관관계 분석(WP → AP)

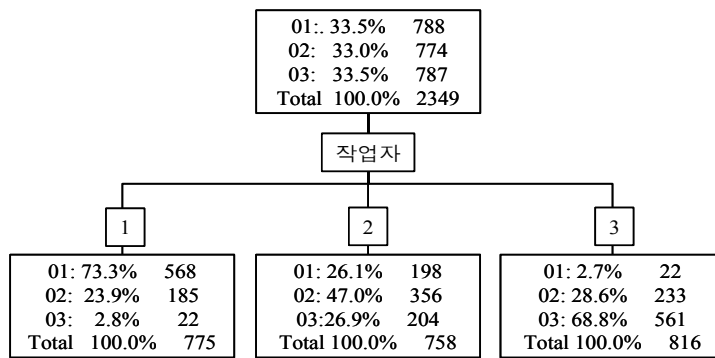
마지막으로 ‘신원조회’ 업무의 수행도 차이의 원인을 파악하기 위한 단계로 작업자의 수행도에 따른 업무의 수행시간의 차이를 분석하였다. 앞서 설명한 바와 같이 수행도가 다른 3명의 작업자에 따라 업무의 수행도의 차이가 구분되었으며 이는 시뮬레이션 모형에서 설정한 사항과 일치함을 다시 알 수 있었다. 그림에서 작업자1이 ‘신원조회’ 업무를 수행했을 때 그 73.3%의 경우에서 해당 업무가

빠르게 완료됨을 알 수 있다. 반면 작업자 3에게 ‘신원조회’ 업무가 할당되었을 경우에는 68.8%의 경우에서 업무가 늦게 완료되었다.

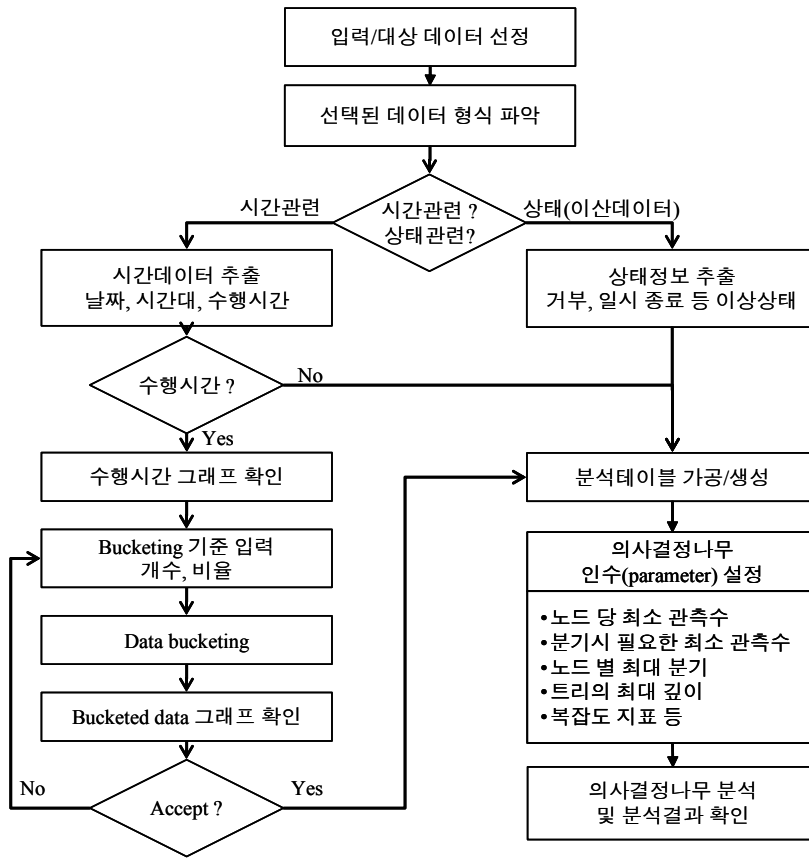
지금까지의 결과를 종합해 보면, 첫째, 보험가입처리 업무의 수행도는 보험가입신청이 물리는 오전에 접수된 경우에서 수행도가 낮은 것으로 파악되었으며, 둘째, 프로세스 흐름상에 병목에 존재하는 ‘신원조회’, ‘온라인 입력’, ‘심사’ 등의 업무에 의해 영향을 많이 받는 것으로 나타났다. 특히 ‘심사’ 업무는 ‘신원조회’ 업무의 수행시간과 높은 상관관계가 있는 것이 밝혀졌다. 마지막으로 ‘신원조회’ 업무의 수행도는 담당작업자의 수행도에 크게 영향을 받는 것으로 나타났다. 도출된 분석결과는 프로세스 상에서 병목이 나타나지 않도록 프로세스를 재설계하거나, 주요한 업무에 수행도가 높은 담당자를 배치시키는 등의 개선작업을 통해 ‘보험가입 처리’ 프로세스의 수행도를 개선시킬 수 있음을 시사한다.

### 4.3 분석시스템

지금까지 설명한 분석절차를 위한 분석시스템을 구축하였다. 분석시스템은 <그림 10>



<그림 9> 작업자에 업무처리 능력에 따른 ‘신원조회’ 업무의 수행시간 차이



〈그림 10〉 분석을 위한 데이터 가공 및 인수 설정

과 같은 절차로 작동한다. 분석시스템의 주요한 기능은 시뮬레이션 모형에서 생성된 BPM 데이터베이스로부터 사용자가 선택한 분석대상 및 입력요인들과 관련된 데이터들을 추출하여 분석에 사용할 테이블 구조로 변환하는 작업이다. 이는 크게 (1) 입력/대상 데이터가 연속형일 경우 필요에 따라 이를 이산형으로 변환하는 작업, (2) 변환된 데이터를 종합하여 분석테이블을 구축하는 작업, 그리고 (3) 의사결정나무 분석을 위한 인수를 설정하는 작업으로 구성된다. 본 연구에서는 BPM 시스템으로 H사의 제품을 참고하였으며, 데이

터마이닝 도구로는 SAS의 e-miner 모듈을 이용하였다.

## 5. 결 론

본 연구에서는 데이터마이닝 기법의 하나인 의사결정나무 분석을 사용하여 비즈니스 프로세스를 분석하는 방법론을 제시하였다. 제시된 방법론은 프로세스, 단위업무, 작업자의 수행도 지표들 간의 연관관계를 체계적으로 추적할 수 있는 방안을 제시한다. 제시된

방법론의 효용성을 입증하기 위하여 시뮬레이션을 통해 여러 가지 가상의 문제 상황을 설정하여 그에 따른 프로세스 이력데이터를 획득하였고, 이를 본 연구의 분석절차를 통해 분석함으로써 제시된 방법론이 설정된 문제를 효과적으로 분석해 낼 수 있는지를 검증하였다.

본 연구에서 제시한 방법론은 비즈니스 프로세스의 수행도 지표의 1차적인 모니터링에 그치지 않고, 이에 영향을 끼치는 요인들을 파악할 수 있게 함으로써 기업으로 하여금 프로세스 개선에 활용할 수 있는 객관적이고 유용한 분석결과를 제공해 준다. 분석을 통해 파악된 문제 요인들에 대한 적절한 통제를 통하여 기업은 직접적으로 비즈니스 프로세스의 효율을 높일 수 있을 것이다. 또한, 몇몇 전문가들의 경험에 의하여 수행되던 프로세스의 재설계, 조직구조의 변화 등의 주요한 혁신 활동을 보다 객관적으로 수행할 수 있는 근거 자료를 본 분석을 통해 획득할 수도 있을 것이다.

본 연구에서는 가상의 프로세스 데이터에 대한 실험을 통해 프로세스 분석방법론의 적용가능성을 간접적으로 확인하였으나, 실제 업무 프로세스의 혁신 과정에 적용되기 위해서는 보다 많은 사례에 대한 적용을 통한 검증이 필요할 것으로 보인다. 또한 본 분석방법이 프로세스 개선에 필요한 객관적 근거를 제공하긴 하지만, 실제적으로 프로세스를 어떻게 재구축해야 하는가에 대한 구체적인 방법을 제시해 주지는 않는다. 분석 결과가 어떻게 프로세스와 조직변화에 활용될 수 있으며 실제로 기대되는 효과를 얻을 수 있는가에 대한 부분은 본 연구에 기반을 둔 다양

한 사례 분석 및 추후 연구의 주요한 과제가 될 것이다.

---

## 참 고 문 헌

---

- [1] Basu, A. and Kumar, A., "Research Commentary : Workflow Management Issues in e-Business," *Information Systems Research*, Vol. 13, No. 1, 2002, pp. 1-14.
- [2] Breiman, L., Friedman, J. H., Olson, R. A., and Kumar, V., *Classification and Regression Trees*, Chapman and Hall/CRC, New York, 2006.
- [3] Davenport, T. H., *Process Innovation*, Harvard Business School Press, 1993.
- [4] Dresner, H., "Business Activity Monitoring : BAM Architecture," Gartner Group; [www.pilos.net/documents/german/gartner.pdf](http://www.pilos.net/documents/german/gartner.pdf).
- [5] Grigon, D., Casati, F., Castellanos, M., Dayal, U., Sayal, M., and Shan, M., "Business Process Intelligence," *Computers in Industry*, Vol. 53, No. 3, 2004, pp. 321-343.
- [6] Hammer, M. and Champy, J., *Reengineering the Corporation*(Revised and update with a new authors' note), Collins Business Essentials, 2005.
- [7] Hur, W., Bae, H., and Kang, S., "Customizable Workflow Monitoring," Con-

- current Engineering-Research and Applications, Vol. 11, No. 4, 2003, pp. 313-326.
- [8] Hwang, Y. and Leitch, R. A., "Balanced Scorecard : Evening the odds of successful BPR," IT Professional, 2005, pp. 24-30.
- [9] Kim, Y., Bae, J., Bae, H., and Kang, S., "Automatic Control of Workflow Processes using ECA Rules," IEEE Transactions on Knowledge and Data Engineering; Vol. 16, No. 8, 2004, pp. 1010-1023.
- [10] Kumar, A. and Zhao JL., "Dynamic routing and operational controls in workflow management systems," Management Science; Vol. 45, No. 2, 1999, pp. 253-277.
- [11] Kung, P., Hagen, C., Rodel, M., and Seifert, S., "Business Process Monitoring and Measurement in a Large Bank : Challenges and selected Approaches," Proceedings of the 16th International Workshop on Database and Expert Systems Applications (DEXA 2005), 2005.
- [12] Srinivasan, S., Krishana, V. and Holmes, S., "Web-Log-Driven Business Activity Monitoring," Computer, Vol. 38, No. 3, 2005, pp. 61-68.
- [13] Thomas, M., Redmond, R., Yoon, V., and Singh, R., "A Semantic Approach To Monitor Business Process Performance," Communications of the ACM, Vol. 48, No. 12, 2005, pp. 55-59.
- [14] Workflow Management Coalition, "Process Definition Interchange," WfMC-TC-1016-P, Lighthouse Point, Florida, 1999.

저 자 소 개



허원창 (E-mail : wchur@inha.ac.kr)  
2004년 서울대학교 산업공학 (박사)  
2005년 ㈜CyberMed 연구소장  
2005년~현재 인하대학교 경영학부 조교수  
관심분야 BPM, 기술혁신, IT entrepreneurship, 인터넷 비즈니스



배혜림 (E-mail : hrbae@pusan.ac.kr)  
2002년 서울대학교 산업공학 (박사)  
2003년 삼성카드, 정보기획팀  
2003년~2004년 동의대학교 인터넷비즈니스학과  
2004년~현재 부산대학교 산업공학과 조교수  
관심분야 BPM, eAI, 웹서비스, 물류 프로세스 관리



김 승 (E-mail : seung2@netopia.snu.ac.kr)  
1999년 부산대학교 (학사)  
1999년~현재 서울대학교 산업공학과 (박사과정)  
관심분야 BPM, Genetic Programming, Web Search



정기성 (E-mail : ksjeong@neoplux.com)  
2001년 서울대학교 산업공학 (석사)  
2005년 딜로이트 컨설팅  
2006년~현재 ㈜네오플렉스 부장  
관심분야 BPM, 비즈니스 컨설팅, 기업투자