건설생산성 관리 시스템 구축을 위한 데이터웨어하우스의 적용

The Application of Data Warehouse for Developing Construction Productivity Management System

오 세 욱* · 김 명 호** · 김 영 석***
Oh, Se-Wook · Kim, Myoung-Ho · Kim, Young-Suk

요 약

건설생산성은 단위작업에서 발생되는 자원의 활동과 이로 인한 산출물간의 관계를 지표로 표현한 것을 의미하며 수집된 건설 생산성 정보는 공사 진행의 효율성 파악, 작업여건 및 투입자원의 성능이 공정에 미치는 영향도 분석, 프로젝트별 성과 측정, 향후 공사계획을 위한 참고자료로서의 활용 등 프로젝트 관리를 위한 다양한 목적으로 이용될 수 있다. 본 연구의 목적은 데이터웨어하우스, OLAP, 데이터 마이닝 기술을 활용하여 PDA 및 바코드를 통해 수집된 생산성 데이터를 축적 및 분석하고, 생산성 예측 모델을 제시하기 위한 방법론을 제안하는 것이다. 이를 통해 보다 효율적이고도 다차원적인 생산성 분석이 가능한 건설 생산성 관리 시스템을 개발하는 것이며, 개발된 시스템은 건설 사업의 합리적인 성과 측정과 유사 프로젝트의 공사계획 수립에 활용할 실적데이터를 축적할 수 있을 것이다.

키워드 : 건설 생산성, 데이터웨어하우스, OLAP, 데이터 마이닝, 생산성 영향요인

1. 서 론

1.1 연구의 배경 및 목적

생산성은 모든 생산조직의 견실성 및 생산 활동의 효율성 평가를 위한 매우 중요한 경영 지표 중 하나로써, 이러한 생산성의 기능은 건설산업과 건설 프로젝트에 있어서도 예외일 수 없다. 건설산업에 있어 생산성 정보는 공사 진행의 효율성 파악, 시공방법, 작업여건 및 투입 자원의 성능이 전체 공정에 미치는 영향도 분석, 프로젝트별 성과 측정, 향후 공사 계획을 위한 기초자료로써의 기능을 지닌다. 그러나 건설산업에서 생산성이 차지하는 중요성과 그 기능에도 불구하고 국내 건설 산업의 경우 생산성 데이터의 수집 및 분석 방법이 체계화되어 있지 못하고 건설 생산성 관리를 대부분 현장관리자의 경험과 직관에 의존하고 있으며, 생산성 데이터를 데이터베이스화하여 공사계획 및 관리에 적극 활용하지 못하고 있는 점은 문제점으로 지적되어 왔다. 최근 업계와 학계를 중심으로 정보기술을 기반으로 한 건설 생산성 데이터의 수집 및 분석에 관한 기초적인 연구들이 시도되어 왔으나 기존 연구의 경우, 건설 생산성에 영향을 미치는 다양한 요인을 고려하지 못하고 다수의 현장에서 발생되는 생산성 데이터를 지속적으로 축적할 수 있는 방법론을 제시하고 있지 못한 점 등은 현재까지도 계속적인 문제점으로 지적되어 왔다. 또한 기존 연구에서 제시한 관계형 데이터베이스 기반의 client/server방식의 시스템은 미리 정해진 양식에 의한 데이터 분석만이 가능하여 생산성 분석의 획일성을 가져올 수 있으며, 분석된 생산성 데이터를 기반으로 한 예측 모델을 효과적으로 제시하지 못함으로써 향후 유사 프로젝트의 공사 계획에 생산성 데이터를 재활용하지 못하는 문제점이 있는 것으로 분석되었다.

오늘날 데이터웨어하우스(data warehouse) 및 OLAP(on-line analysis process)기술은 이와 같은 운영계 시스템에서의 데이터 축적 및 분석의 문제점을 해결하기 위해 등장하였으며 이미 타 산업 전반에 적용되고 있다. 또한 데이터의 축적 및 분석에만 그치지 않고 축적된 데이터를 활용하여 새로운 지식을 발견하는 데이터 마이닝(data mining) 기술 또한 다양한 분야에 적용되고 있다. 따라서, 데이터웨어하우스 및 OLAP, 데이터 마이닝 기술을 건설 산업의 생산성 관리에 접목한다면 생산성

* 일반회원, 인하대학교 건축학부 Postdoctoral Fellow, 공학박사, archiom@inha.ac.kr
** 일반회원, 인하대학교 건축학부, 공학석사, yasha78@korea.com
*** 종신회원, 인하대학교 건축학부 부교수, 공학박사(교신저자), youngsuk@inha.ac.kr

데이터는 공사계획 및 관리측면에서 보다 효율적이고도 다양하게 활용될 수 있을 것으로 사료된다.

본 연구에서는 데이터웨어하우스, OLAP, 데이터마이닝 기술을 수집된 생산성 영향요인 및 생산성 데이터에 적용하고 이를 통해 건설 생산성 데이터의 축적 및 분석과 예측 모델을 제시하기 위한 방법론을 제안하고자 한다. 이는 보다 효율적이고 다차원적인 생산성 분석이 가능한 건설 생산성 관리 시스템을 개발하는 것을 의미하며 개발 시스템이 건설 현장에 적극 활용될 경우, 건설 생산성 분석의 유연성 제고를 통한 생산성 관리, 합리적인 건설 사업의 성과 측정뿐만 아니라 향후 유사 프로젝트에 있어 생산성 예측을 통한 보다 신뢰성 있는 공사계획 수립 등이 가능할 것이다.

1.2 연구의 범위 및 방법

건설 생산성의 효율적인 축적과 다양한 분석을 위해서는 생산성에 영향을 미치는 다양한 요인을 도출하고 해당 요인을 반영하여 생산성 데이터의 축적 및 분석이 이루어져야 한다. 본 연구에서는 공동주택 구체공사를 대상으로 기존 문헌 고찰 및 현장 관리자와의 인터뷰를 통해 건설 생산성에 영향을 미치는 정량적 · 정성적 요인을 도출하고, PDA, 바코드, 데이터웨어하우스, OLAP, 데이터마이닝 등의 정보기술을 활용하여 수집된 생산성 데이터를 관리자가 다양한 관점에서 분석 및 예측할 수 있는 건설 생산성 관리 시스템을 제시하고자 한다.

이를 위한 연구의 방법은 다음과 같다.

1) 국내 · 외 연구 동향 분석 ; 건설 생산성 관리와 관련된 국내 · 외 연구 동향을 고찰하여, 생산성 데이터의 분석 및 시스템 구축 측면에서 기존연구의 한계성을 도출하였다.

2) 건설 생산성 영향요인 도출; 본 연구에서는 먼저 건설 생산성의 개념을 간략히 정의하고, 기존 연구문헌 고찰과 현장 인터뷰를 통해 건설 생산성에 영향을 미치는 다양한 요인을 조사 및 분석하여 공동주택 구체공사를 대상으로 한 건설 생산성 영향요인을 도출하였다.

3) 생산성 데이터웨어하우스 개발; 도출된 생산성 영향요인을 반영하고 다양한 건설 프로젝트에서 발생하는 생산성 데이터를 효과적이고도 신속하게 축적 및 분석하기 위한 데이터웨어하우스를 설계 및 구축하였다.

4) 건설 생산성 데이터의 다차원적 분석을 위한 OLAP 시스템 개발; 수집된 건설 생산성 데이터를 다차원 정보로 가공하기 위해 OLAP 서버를 구축하고 관리자가 가공된 다차원 정보에 접근하여 유연하고 다양한 관점에서 데이터를 분석할 수 있도록 하는 OLAP 클라이언트를 Visual Basic 6.0을 이용하여 개발하였다.

5) 데이터 마이닝 기술을 이용한 건설 생산성 예측 모델 제시; 축적된 건설 생산성 데이터를 향후 유사 프로젝트의 공사계획 시 활용하기 위하여 데이터 마이닝 기법 중 신경망 알고리즘을 이용한 건설 생산성 예측 모델을 제시하였다.

6) 개발 시스템의 검증; 선행 연구(오세욱, 2005)를 통해 공동주택 구체공사를 대상으로 수집된 생산성 데이터를 적용하여 건설 생산성 분석 및 예측을 위한 개발 시스템의 효용성을 검증하였다.

2. 건설 생산성 관련 국내 · 외 연구동향 분석

건설 생산성 데이터의 분석 및 예측과 관련된 기존 연구 동향을 시스템 구축측면과 생산성 데이터의 분석 및 예측 측면으로 구분해서 살펴보면 다음 표 1과 같다.

표 1에서 제시된 선행연구의 한계성은 다음과 같이 요약정리될 수 있다.

1) 건설 생산성에 직 · 간접적으로 영향을 미치는 주요 요인들을 도출하는 과정은 일부 기술하고 있으나 주요 요인들에 대한 객관적인 데이터의 반영과 이에 따른 생산성 저하 및 향상에 대한 원인 규명, 대처 방안에 대한 기술은 미흡하다.

표 1. 선행 연구에 대한 이론적 고찰

구분	연구자	내 용
시스템 구축	오세욱 (2005)	공동주택 구체공사에 있어 단위작업을 기반으로 노무자의 작업 시작 및 종료시간, 진행물량 정보를 PDA 및 바코드로 수집하고 이를 토대로 주요 생산성 데이터를 도출하는 전산화 시스템을 구축하였다. 2개의 현장 적용을 토대로 생산성 데이터가 비교되어 학습 곡선, 작업시기, 작업 난이도 및 작업조의 숙련도 등이 공동주택 골조공사의 생산성에 영향을 주는 것으로 분석되었음
	Carlibre 2000(BRE,2000)	현장관리자가 휴대용 단말기 및 디지털 카메라를 이용하여 단위작업 투입된 노무자 작업시간 및 비 작업시간, 휴식시간 등의 정보와 실적물량 정보를 수집하여 이를 데이터베이스화하고 생산성 추이를 각종 그래프 및 차트를 활용하여 표현함으로써 생산성 검측을 통한 성과측정의 결과를 분석할 수 있도록 함.
	유정호 (2002)	건설 생산 단계에서의 생산성관리를 통한 생산성 향상 및 엔지니어링 과정의 효율성 향상을 위해 시공과정에서의 생산성 데이터를 체계적이고 지속적으로 수집 및 가공하고 관리 할 수 있는 방법론과 생산성 데이터의 관리를 위한 데이터 모델을 제시하였다. 이를 통해 생산성 데이터를 시공 과정에서의 생산성관리에 활용하는 방안과 엔지니어링 과정에 활용하는 방안을 제시함
생산성 데이터 분석 · 예측	안용선 (1993)	3개 공동주택 현장 거푸집 공사를 대상으로 단위작업에 투입된 노무자의 작업시간 및 실적물량을 연속시간 측정 기법으로 수집하고, 부위 공법별로 생산성 데이터 값을 도출함. 또한 3개 현장에 대한 부위 및 공법별로 생산성 데이터 값의 차이를 분석하였으며 이를 토대로 소요 품에 대한 비용 산정 방안을 제안함
	Rifat sonmez (1995)	건설 생산성에 주요 영향요인으로 수량, 기온, 습도, 작업조 크기, 시간외 작업 등으로 구분하고 도출된 영향 요인별로 건설생산성에 효과를 평가하는 방안을 신경망 알고리즘을 통해 제안함
	표영민 (2005)	노동 생산성을 저하시키는 정성적인 영향 인자들을 도출하고 영향 인자들 간에 영향도와 우선 순위를 AHP 기법을 통해 분석하여 저하 요인별 생대적인 중요도를 산출 함

2) 대부분의 기존 연구들은 생산성 수집이나 향상 방안을 중점적으로 다루고 있을 뿐 생산성 분석을 통한 성과측정 및 축적된 데이터의 구체적인 활용방안을 제시하지 못하고 있다.

3) 생산성 데이터의 분석 작업은 많은 데이터베에스 자원을 필요로 하기 때문에 시스템 구축 측면을 단일 현장 중심으로 구축하는 것에는 한계성이 있다.

4) 시스템 구축에 있어 관계형 데이터베에스의 클라이언트/서버 방식은 개발자가 미리 정해놓은 툴(tool)에 맞춰 분석하기 때문에 생산성 분석에 있어 획일성을 초래할 우려가 있다. 이러한 획일적인 시스템 구조는 원하는 정보를 적시에 획득할 수 없어 변수가 많은 건설 생산성 데이터 취합을 위한 시스템 구조로는 적합하지 않다.

따라서, 본 연구에서는 건설 생산성 데이터와 영향요인간의 관계를 보다 객관적이고도 실증적인 데이터로서 분석할 수 있도록 하고 여러 현장에서 축적된 생산성 데이터를 사용자 중심에서 다양한 각도로 분석 및 예측할 수 있도록 하는 방안을 제안하고자 하며, 데이터웨어하우스에서의 OLAP, 데이터마이닝, 신경망 알고리즘 모델을 적용한 건설 생산성 관리시스템을 제안하고자 한다.

3. 건설 생산성 관리 체계

3.1 건설 생산성

생산성은 투입된 자원(input)에 따른 산출물(output)의 비로써 생산체계의 효율성을 측정할 수 있는 지표로 정의할 수 있으며, 이때 투입 자원과 산출물을 무엇으로 규정할 것인가에 따라 다양한 측정 방법이 활용될 수 있다. 건설 생산성의 경우에도 산업차원의 거시적 수준에서부터 공종 또는 세부 수준의 단위작업에 이르기까지 다양한 정의가 가능하나 프로젝트의 효율적 관리와 시공능력 향상, 그리고 생산성 데이터의 기능적 활용을 위해서는 공종 및 단위작업의 생산성 자료의 수집과 분석이 필수적이다. 공종 및 단위작업 수준의 생산성 측정에서 투입된 자원은 노무량으로 정의될 수 있으며 산출물은 진행된 작업 물량으로 정의된다. 본 연구에서의 건설 생산성은 단위작업을 기반으로 발생된 작업정보인 노무 작업시간과 진행 물량간의 비(Qty/M · H, M · H/Qty)로 정의되며, 단위작업을 기반으로 한 생산성 자료의 활용 예를 살펴보면 다음 그림 1과 같다.

3.2 건설 생산성 영향 요인

앞서 정의한 건설 생산성 데이터(Qty/M · H, M · H/Qty)들은 다양한 요인들에 의해 직 · 간접적으로 영향을 받는다. 실례로 공동주택 구체공사에 있어 단위작업을 기준으로 발생되는 생산성 데이터는 일반적으로 평형, 작업시기, 작업 난이도, 건물구조, 작업조의 작업능력 등 다양한 요인들에 의해 영향을 받는다. 본 연구에서는 건설 생산성에 영향을 미치는 주요 요인을 도출하기 위해 관련 문헌(김예상 1994, 손창백 2002, 오세욱 2005, Sonmez 1998) 고찰 및 공동주택 골조공사를 중심으로 한 현장 인터뷰를 실시하였으며, 도출된 생산성 영향 요인의 상호 비교 · 분석을 통한 필터링 과정(김명호, 2005)을 통해 건설 생산성에 영향을 미치는 주요 요인을 표 2와 같은 정성적, 정량적 요인으로 분류하였다.

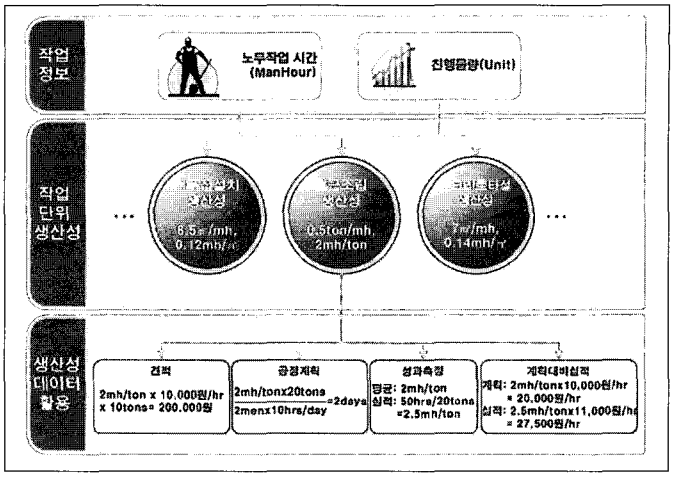

그림 1. 단위작업 기반의 건설생산성 자료 활용

표 2에서 '정량적 요인'은 관리자의 주관적 판단에 영향을 받지 않고 프로젝트의 특성이나 기후조건 등과 같이 객관적으로 그 영향도가 가시적(tangible)이라 판단되는 인자들을 의미한

표 2. 생산성 영향요인 분류

정성적 요인		정량적 요인
분 류	영향요인	영향요인
건설인력	- 기능인력 부족 - 열악한 작업환경 - 작업자의 숙련도 부족	- 평형타입 - 층별 세대수 - 동 형태 - 층 그룹 - 공종별 그룹 - 작업조 구성 - 온도 - 습도 - 강우량 - 공법 - 지역
설계부분	- 설계도서 미흡 - 설계난이도 - 과다한 설계변경	
공사관리	- 작업조간의 간섭 - 의사결정 지연 - 작업지시 및 승인지연 - 안전사고 발생 - 불합리한 공정계획	
자원관계	- 자재조달 지연 - 장비조달 지연 - 부적합한 자재 투입 - 장비의 성능부족	
공사환경	- 파업 및 민원발생	

다. 반면, '정성적 요인'은 관리자의 주관적 판단이나 특정 사건의 발생 정도에 따라 영향도가 인지되는 비가시적인(intangible) 인자를 의미하며, 개발 시스템에서는 생산성 데이터의 분석 시 이러한 정성적 요인을 '생산성 조건(저하 요인)'으로 반영하였으며, 관리자가 주관적 판단에 의해 필요시 그 결과를 PDA를 통해 수집할 수 있도록 설계하였다.

4. 건설 생산성 관리 시스템의 개발

본 연구에서는 단위작업에서 발생된 작업정보를 생산성 데이터로 가공하고 축적하기 위해 데이터웨어하우스를, 도출된 생산성 데이터와 생산성 영향요소간의 관계를 다차원적으로 분석하기 위한 도구로서 OLAP 기술을 적용하였다. 또한 향후 유사 프로젝트의 공사계획 수립을 위한 생산성 예측 값(reference data)을 제시하기 위하여 데이터 마이닝 기술 중 신경망 기법을 활용하였다. 본 연구에서 제시하고 있는 건설 생산성 관리시스템의 개념은 다음 그림 2와 같다.

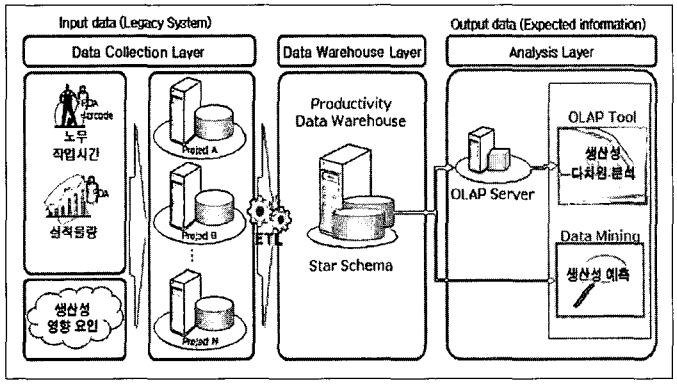

그림 2. 건설 생산성 관리 시스템의 개념

4.1 데이터웨어하우스의 개념 및 구조

Inmon(1992)은 데이터웨어하우스에 대해 "수년간 운영계 시스템에서 발생된 데이터를 주제별로 통합하여 여러 측면으로 분석을 가능케 하는 의사결정 지원 통합시스템"으로 정의하고 있다. 본 연구에서의 데이터웨어하우스는 건설 생산성 데이터의 수집을 통해 발생된 데이터를 생산성이란 주제로 통합하여 생산성 영향요소 측면에서 다양하게 분석하는 것을 의미한다. 또한 데이터웨어하우스 저장소는 관계형 데이터베이스와는 달리 다차원 모델 구조로 표현된다. 다차원모델은 그림 3과 같이 차원과 변수 항목으로 구성되어 있는 입방체(cube)로 이해할 수 있다. 여기서 차원이란 관리자가 정보를 분석하고자 하는 하나의 관점을 나타내며 변수는 수치 또는 측정치로 나타내는 정보를 의미한다. 또한 차원은 일반적으로 계층적 구조를 갖게 되며 계층 구조는 부모(parent)와 자식(child)간의 관계로 존재하게 된다. 즉 부모는 계층적인 구조상에서 어떤 항목의 바로 상위 항목을 나타내며 자식은 하위 항목을 나타낸다(조재희, 1996). 실례로 건설생산성에 있어 평형타입을 부모로서 차원이라 하면 해당 평형대인 20평형, 30평형, 40평형 등을 자식으로 하위 항목이라고 정의할 수 있다. 따라서 건설 생산성에 대한 다차원 모델은 그림 3에서와 같이 평형타입, 층 그룹, 시설물 등을 차원으로 설정하고, 각 차원에 해당되는 세부 하위 항목에 따라 변수 값을 지정할 수 있다. 여기서 변수 값은 건설 생산성의 값(data)을 의미한다.

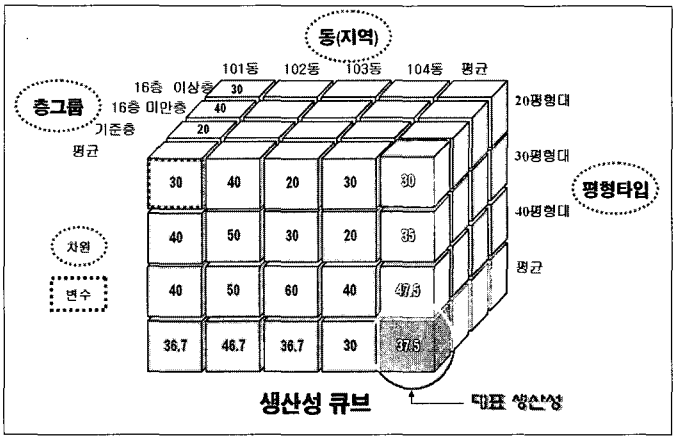

그림 3. 다차원 모델 구조의 예

본 연구에서는 건설 생산성 데이터웨어하우스의 설계를 위해 앞서 정의한 생산성 영향 요인들을 차원으로 두고 각 차원마다 해당 하위 항목을 표 3과 같이 설정하였으며 하위 항목 수준에서 생산성 데이터를 변수로 지정하여 정보를 수집할 수 있도록 하였다.

표 3. 생산성 분석 모델 항목

구분	상위항목	하위항목
변수	측정치	생산성, 투입노무량, 작업물량, 투입노무비
차원	기간	1월, 2월, 3월, ...
	협력업체	A협력업체, B협력업체, ...
	평형타입	28평형, 32평형, 39평형, ...
	층그룹	기준층, 16층 미만층, 16층 이상층
차원	공종	먹메김, 갱폼인양, Wall 거푸집, Wall 철근배근, ...
	동형태	사각형, ㄱ자형, 층계형, ...
	작업조구성	0%~100%
	연장작업비율	0%~100%
	습도	0%~100%
	온도	-20℃~40℃
	강우량	0mm~200mm
	공법	갱폼, 유로폼, 알루미늄 폼, CPB타설, 펌프카타설, ...
	지역	서울, 경기, 인천, ...
	생산성조건	기능인력의 부족, 조악한 작업환경, ...

이와 같이 건설 생산성과 관련된 차원, 각 차원의 하위 항목 구조, 변수를 기반으로 그림 4와 같은 생산성 데이터 모델을 설계하였다. 여기서 변수항목은 팩트 테이블로 표현되며 차원항목은 차원 테이블로 표현된다. 차원 테이블의 기본키(primary key)는 팩트 테이블의 외래 키(foreign key)와 연결되어 관계를 맺는다. 이러한 다차원 데이터 모델링 기법은 중앙의 팩트 테이블과 그 주위를 둘러싸고 있는 차원 테이블의 형태를 취하기 때문에 스타 스키마(star schema)라고 정의된다.

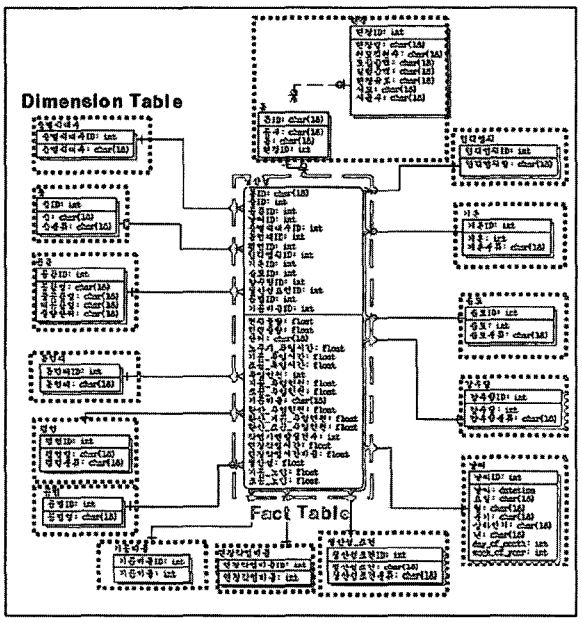

그림 4. 생산성 모델 스타스키마

4.2 데이터 추출, 변환, 적재 프로세스

ETL(extracting, transformation, loading)은 다량의 상이한 데이터를 통합하고 처리할 필요성이 있는 경우, 다양한 DBMS(database management system)와 어플리케이션들로부터 필요한 원천 데이터를 추출하고 정제, 변환하여 대상 데이터베이스에 이동 적재하는 기능으로서 데이터웨어하우스 구성요소간의 일관성과 통합성을 유지시키는 일종의 데이터 통합 엔진을 의미한다(www.hostechglobal.com).

본 연구에서는 기존에 개발된 단위작업 정보 기반의 공동주택 공정관리 지원 시스템인 ASIMO(오세욱, 2005)에서 수집된 현장정보에서 생산성과 관련된 데이터를 추출(extracting)하였다. 또한 기존 시스템에 반영되어 있지 않은 기후 정보 및 작업조 구성, 연장작업 비율은 각각 기상청 자료 및 내부 연산을 통해 데이터를 생성(transformation)하였고 앞서 추출된 생산성 데이터와 함께 하나의 로우셋(row set)을 생성하여 데이터웨어하우스에 적재(loading)하였다. 그림 5는 생산성 데이터의 ETL과정을 도식화 한 것이다.

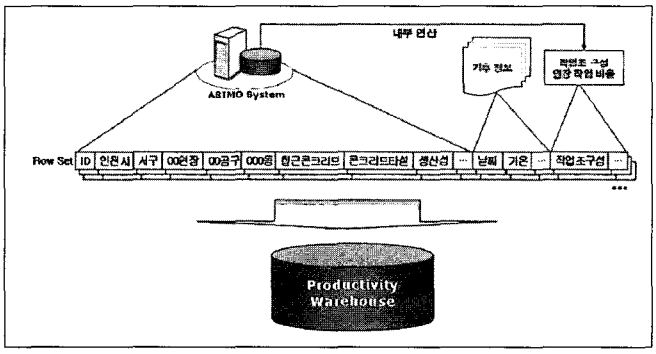

그림 5. 생산성 관련 데이터의 추출, 변환, 적재 프로세스

4.3 다차원 분석을 위한 OLAP시스템 구축

OLAP은 관리자가 다차원 정보에 직접 접근하여 대화식으로 정보를 분석하고 의사결정에 활용하는 과정을 의미한다. ETL알고리즘을 통해 데이터웨어하우스에 축적된 생산성 데이터는 OLAP서버를 통해 다차원 정보로 가공되며 관리자는 OLAP클라이언트를 통해 가공된 다차원 정보에 직접 접근할 수 있다. 본 연구에서는 건설 생산성 분석 시스템의 OLAP서버를 마이크로소프트사의 Analysis Service를 이용하여 14개의 차원 테이블(생산성 영향요소)과 1개의 팩트 테이블(생산성 값)로 구성된 생산성 분석 OLAP서버를 다음 그림 6과 같이 구축하였으며 Visual Basic 6.0을 이용하여 OLAP클라이언트를 그림 7과 같이 개발하였다.

기존의 관계형 데이터베이스 기반의 클라이언트/서버(client/server) 방식을 사용한 분석 방법은 관리자의 분석 관점에 따라 여러 개의 분석화면을 개발해야 하기 때문에 많은 시간과 노력이 필요하다. 그러나 OLAP방식의 다차원 분석 방법은 그림 7에서와 같이 관리자의 다양한 관점 및 자유로운 선택에 따라 다차원적인 데이터 분석이 가능하기 때문에, 기존의 관계형 데이터 분석 방식에 비해 매우 효과적이고도 유용한 분석 툴로 정의될 수 있다.

개발 시스템에서는 그림 7의 (1)에서와 같이 좌측에는 생산성 영향요인 목록이, 우측에는 (2)피벗테이블이 구성되어 있어 관리자가 자신의 분석 관점에 따라 영향 요인 목록을 드래그 앤 드

1) 팩트테이블(fact table): 숫자 측정값 및 팩트와 차원 테이블을 연결하는 키를 포함하는 데이터웨어하우스 스키마의 중앙 테이블로 정의되며 팩트테이블에는 빈 트랜잭션이나 제품 매출 등 업무 관련 특정 이벤트를 설명하는 데이터가 포함 된다(SQL Server 2000 온라인 설명서).

2) 차원테이블(dimension table): 팩트 테이블에 구성된 데이터를 설명하는 데이터웨어하우스 테이블(SQL Server 2000 온라인 설명서).

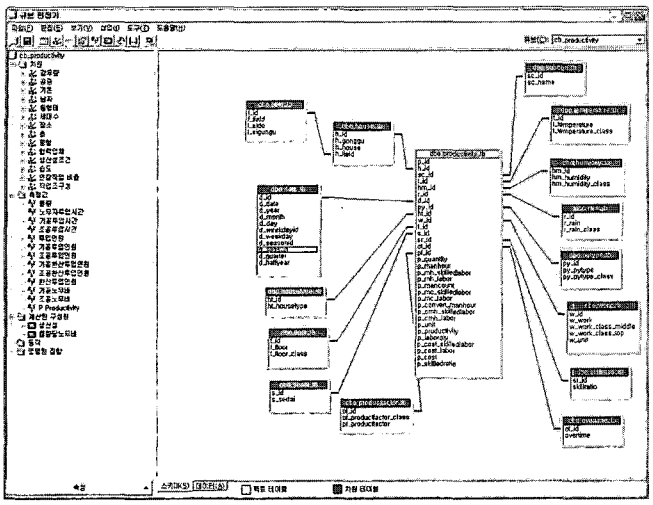

그림 6. OLAP서버 구축 결과

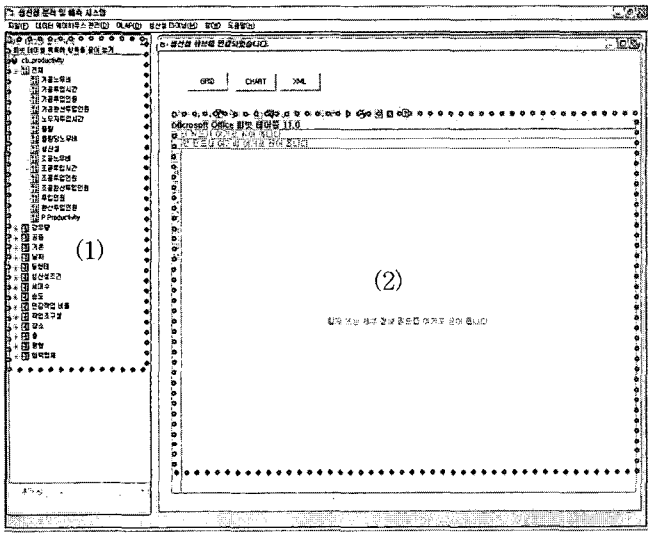

그림 7. OLAP 클라이언트 화면

롭 방식으로 피벗 테이블로 이동시키면 그에 따른 다양한 생산성 데이터가 자동으로 취합되는 구조를 지니고 있다.

4.4 생산성 예측모델의 제시를 위한 인공 신경망 기법의 활용

데이터 마이닝이란 많은 양의 원시 데이터에서 일정한 패턴을 추출하는 특별한 알고리즘으로 구성된 응용기술을 통해 유용한 지식을 도출해내는 기술을 의미한다(Han, 2001). 건설 생산성 예측은 다양한 인자(생산성 영향요인)가 상호 복합적인 영향을 받아 정량적 데이터로 표현하는 것을 의미하기 때문에 데이터의 일정한 패턴을 추출하는 과정이 필요하다. 본 연구에서는 데이

3) 피벗 테이블(pivot table) : 데이터를 요약하고 분석할 수 있도록 해주는 대화형 테이블 컴포넌트(component)

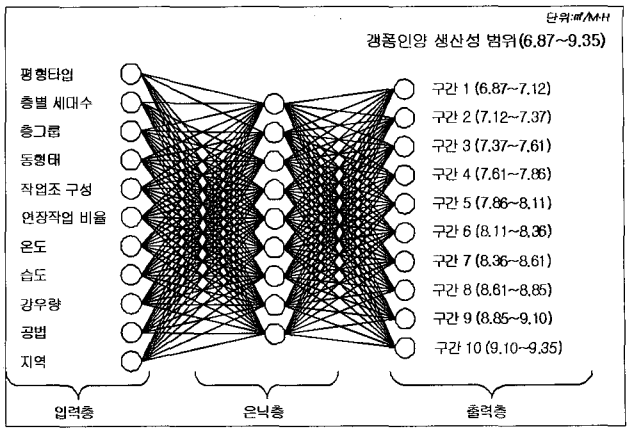

그림 8. 건설 생산성 예측을 위한 인공 신경망 모델

터마이닝을 적용하여 건설 생산성 예측 모델을 제시하고자 하며 데이터 마이닝의 다양한 기법 중 데이터 예측에 우수한 성능을 가진 인공 신경망(artificial neural network) 기법을 활용하고자 한다. 인공 신경망은 인간의 뇌 구조의 형태를 통해 학습능력을 가진 수리모델을 모델링 하는 것으로 매우 복잡한 구조를 가진 데이터 사이의 관계나 패턴을 찾아내는데 유용한 비선형 모형이다. 인공 신경망 모델의 입력 층은 생산성 영향 요인 중 정량적 요인인 평형타입, 층별 세대수, 강우량, 공법, 지역 등 총 11개의 유닛(unit)으로 이루어져 있으며 은닉 층의 유닛 수는 관리자가 직접 선택할 수 있게 하였다. 또한 출력 층은 단위작업별로 수집된 생산성 데이터의 최대치와 최소치를 10개의 구간 나누어 설정하였다. 그림 8은 생산성 예측과 관련한 인공 신경망 모델을 도식화 한 것이다.

인공 신경망 모델을 통해 계산된 출력 값과 실제 데이터간의 오차를 줄이는 과정은 신경망의 반복적 학습을 통해 이루어지며 본 연구에서는 신경망 학습을 위해 백프로퍼게이션(backpropagation)알고리즘을 사용하였다. 백프로퍼게이션 학습 알고리즘은 입력 층의 각 유닛에 입력 패턴 신호를 주면, 이 신호에 따라 각 유닛에서 입력 패턴 내용이 변환되어 중간층에 전달되고 최후에 출력 층에서 신호에 따른 출력 값과 기대 값을 비교하여 차이를 줄여나가는 방향으로 연결 강도를 조정해 가는 학습방법이다(김대수, 1992). 이와 같은 백프로게이션 알고리즘을 이용한 건설 생산성 예측 모델의 데이터 처리과정은 다음 그림 9와 같다. 그림 9에서와 같이 건설 생산성 예측 모델의 데이터처리과정은 크게 데이터 준비단계, 신경망 학습단계, 신경망 모델의 저장 및 예측단계로 구분할 수 있다. 먼저 데이터 준비단계에서는 신경망 학습을 위해 데이터웨어하우스에 축적되어 있는 생산성 데이터의 수치를 추출하고 텍스트(text)형태의 속성을 갖고 있는 생산성 영향요인은 숫자로 변환하여 데이터베이스에 저장한다. 또한 신경망 알고리즘의 출력 층에서 설계한 10개

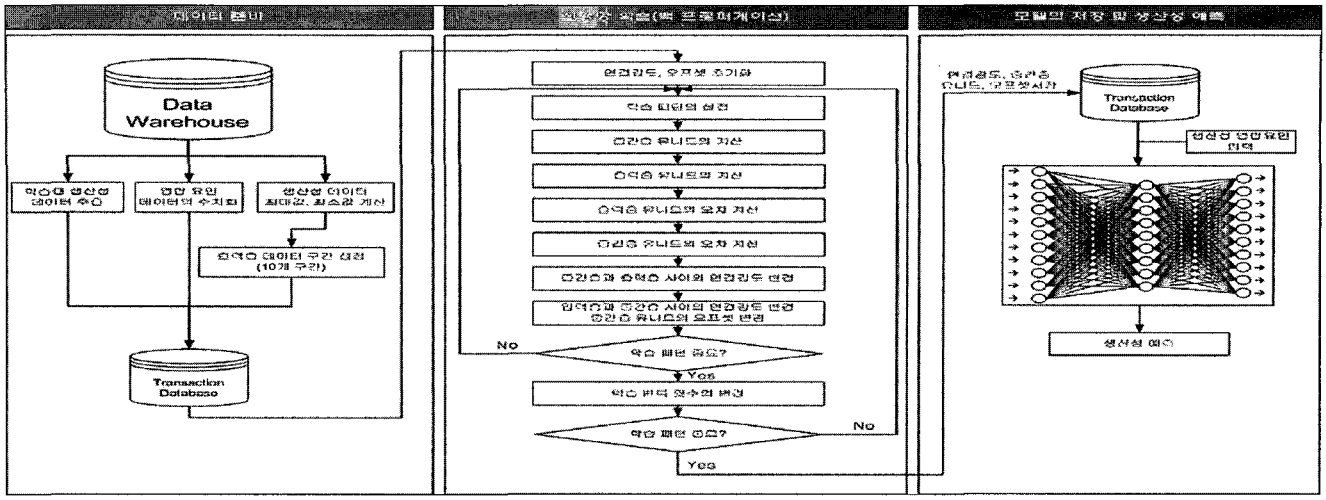

그림 9. 건설 생산성 예측 모델의 데이터 처리과정

의 생산성 데이터 구간 범위를 설정하여 데이터베이스에 저장한다. 신경망 학습단계에서는 앞서 설명한 백프로퍼게이션 알고리즘 통해 학습되며 학습된 결과에 따라 연결강도, 중간층 유닛 값, 오프셋 작업별로 데이터베이스에 저장된다. 신경망 모델의 저장 및 예측단계에서는 이와 같은 절차를 통해, 관리자가그림 . 건설 생산성 예측 모델의 데이터 처리과정 지정한 영향요인에 따라 적정한 생산성 데이터 수치를 예측할 수 있게 된다.

본 연구에서 제시하고 있는 생산성 예측 시스템은 Visual Basic 6.0으로 구현하였으며 생산성 모델을 저장할 데이터베이스는 MS-SQL Server 2000으로 구현하였다.

5. 개발 시스템의 효용성 검증

본 연구에서는 선행 연구(오세욱, 2005)를 통해 수집된 2개 공동주택 현장의 작업 정보 중, 아래 표 4와 같은 1개 현장에서 수집된 생산성 데이터를 토대로 건설 생산성 데이터의 분석 및 예측을 위한 개발 시스템의 효용성을 시험적으로 검증하였다.

표 4. 사례 현장의 정보 수집 개요

공사명	인천 A아파트 신축공사
세대수/층수	11개동(719세대), 지하 2층, 지상 20층
구조방식	철근콘크리트 벽식구조
공사진행	정보 수집 시점에서 각 동의 13층부터 14층 구체공사를 진행 중(공정률 38퍼센트)
대상 범위	1공구 5개동 5개층
대상단위작업	총 9개 : 먹메김, 옹벽철근조립, 슬라브철근조립, 콘크리트 타설, 거푸집탈형, 슬라브거푸집설치, 옹벽거푸집설치, 갱폼인양, 계단부 거푸집 설치
주요 정보수집	단위작업에 투입된 노무자 작업시간 및 실적물량, 작업 생산성 저하 요인
정보수집 기간	2개월(60여일)

참고적으로, 사례 현장에 있어 기 수집된 생산성 데이터(Quantity/M · H)의 경우에는 단위작업별 생산성 데이터 값과 생산성 저하요인, 평형별, 층 그룹별, 동별 세대수 등의 생산성 영향요인을 일부 포함하고 있다. 또한 본 연구의 적용성 검증을 위해 수집되지 않은 생산성 영향요인들은 정보수집 기간 동안 작성된 작업일보와 날씨, 협력업체 정보 등을 역으로 추적하여 그 결과를 생산성 영향 요인에 추가로 반영함으로써 개발 시스템의 효용성을 검증하였다. 이와 같은 결과를 토대로 개발 시스템이 제공하는 몇가지 주요 화면을 간략히 소개하고자 하며 소개된 화면 뿐 만아니라 표 3에서 제시한 다양한 차원별로 변수(생산성 데이터)를 표현할 수 있다.

1) 생산성 추이 분석

그림 10은 slab 철근 배근과 wall 철근 배근에 대한 기간별 생산성 변화의 추이(trend)를 보여주는 것으로 (1)은 10월 27일 slab 철근 배근에 대한 작업이 진행되었고 그날의 생산성이 0.36tf/M · H임을 알 수 있으며, 28일에는 wall 철근 배근에 대한 작업이 진행되었고 그날의 생산성이 0.43tf/M · H임을 파악할 수 있다. (2)는 월별 생산성의 대표 값을 보여주고 있는 것으로 10월 1개월 동안 slab 철근 배근에 대한 생산성 값이 0.32tf/M · H이고 wall 철근 배근에 대한 생산성 값이 0.37tf/M · H임을 파악할 수 있다. (3)은 연도별 생산성 대표 값을 의미하며 slab 철근 배근의 경우 0.3tf/M · H이고 wall 철근 배근의 경우 0.39tf/M · H임을 파악할 수 있다. 또한(4)에서와 같이 가로축에는 날짜를 세로축에는 생산성 수치를 기준으로 하여 slab 철근 및 wall 철근 배근 등에 대한 작업진행 현황을 꺽은 선 그래프의 형식으로 표현함으로써 관리자가 분석하고자 하는 해당 기간(일/주/월/년)의 단위작업별 생산성 추이의 변화를

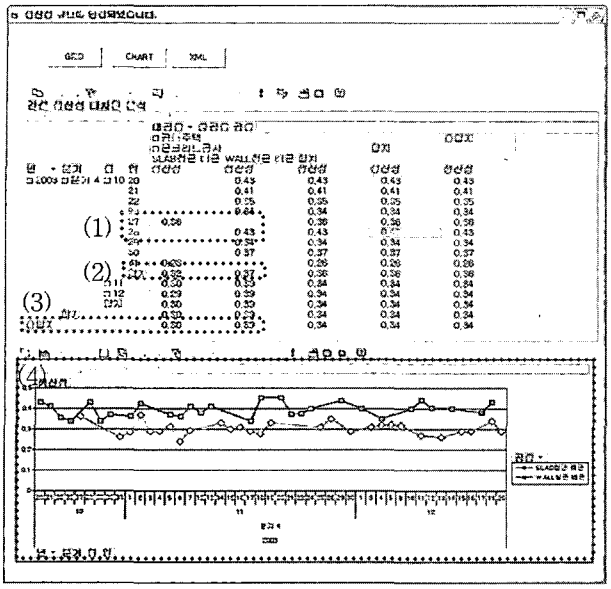

그림 10. 기간별 생산성 분석

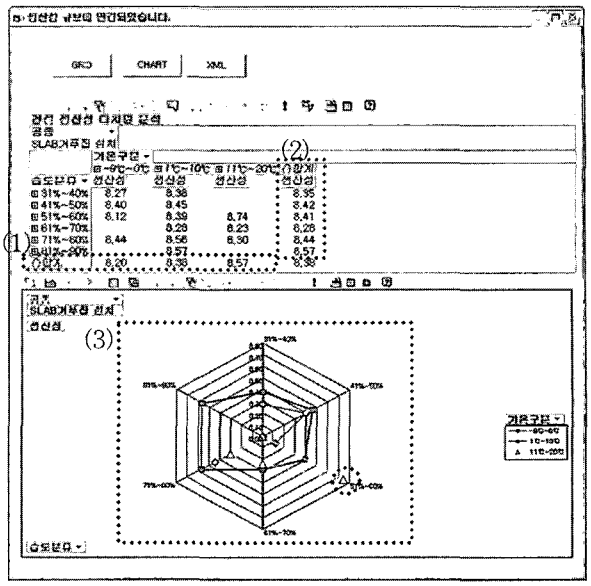

그림 12. 온도 ℃및 습도별 생산성 분석

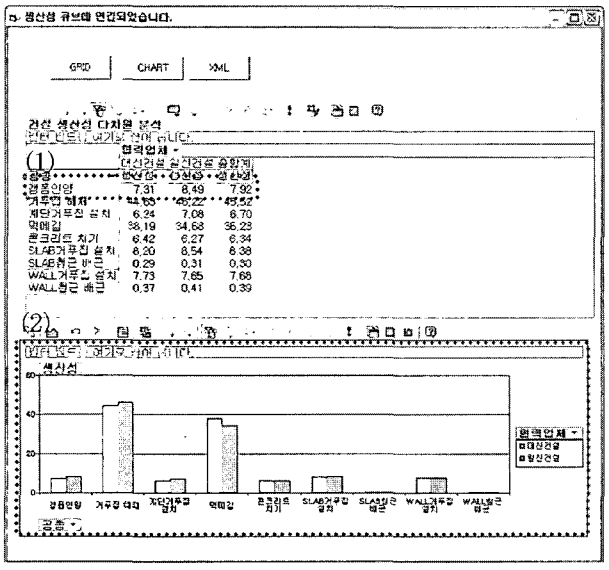

그림 11. 협력업체별 생산성 분석

매우 용이하게 파악할 수 있도록 하였다.

2) 협력업체의 성능 분석

그림 11은 협력업체별로 생산성을 분석하는 화면으로서 사례 현장의 경우 2개의 철근 콘크리트 공사 전문업체가 투입되었으며, 개발 시스템의 경우 프로그램 내에서 기 그룹핑된 동일 혹은 유사조건 하에서 작업을 수행하는 협력업체별 단위작업의 생산성 값을 상호 비교?분석할 수 있다. 실례로 (1)에서와 같이 갱폼 인양의 경우 D업체는 공사개시 이후 현재에 이르기까지(to-date) 7.31M2/M · H를 수행하고 있는데 반해 I업체의 경우에는 8.49M2/M · H를 수행함으로써 I 업체가 D업체에 비해 상대적으로 우수한 성능을 발휘하고 있음을 알 수 있다.

또한 (2)는 가로축에 단위작업들을 세로축에는 생산성 수치로 설정하고 각 작업마다 협력업체별 생산성 값을 막대 그래프형식으로 표현하고 있다. 이러한 협력업체별 생산성 분석 결과는 협력업체의 성능 측정 및 작업 촉진(expediting)을 위한 객관적인 자료를 제공할 수 있으며 협력업체간의 생산성을 비교 및 분석함으로써 협력업체에 대한 생산성 향상의 자발적 동기부여를 제공하여 기존 방식에 비해 보다 효과적인 업체관리를 가능케 할 수 있다.

3) 기후특성 및 생산성 조건에 따른 영향도 분석

그림 12는 slab 거푸집 설치 작업에 있어 기온 및 습도별 생산성 분석 결과를 표현하고 있는 화면으로 가로축에는 온도와 관련된 사항을 세로축엔 습도와 관련된 사항을 나열하고 각 항목에 해당 되는 수치를 표현하고 있다. (1)은 온도를 10℃ 단위로 그룹핑하여 생산성 값을 표현하고 있으며 (2)는 습도를 10퍼센트 단위로 그룹핑하여 생산성 값을 표현하고 있다. 또한 (3)은 극좌표 그래프를 통해 도출된 생산성 값을 표현하고 있는 화면으로 slab 거푸집 설치 작업의 경우, 습도가 60~70퍼센트이고 온도가 11℃에서 20도℃일 경우에 생산성이 가장 높은 것으로 분석되었다.

그림 13은 표 2의 정성적 영향 요인들에 의한 생산성 분석 화면으로 특정 공종(콘크리트 치기)에 대해 가로축에는 시설물 정보를 세로축에는 생산성 조건 차원에 따른 정성적 요인들을 두고 각 항목에 해당되는 생산성 값을 표현하고 있다. 이는 현장관

리자가 해당 일의 투입노무자 및 실적물량 정보를 수집한 시, 특정 공종에 있어 생산성 저하가 발생하였다고 판단하였을 경우 입력된 내용을 토대로 표현된 결과물을 의미한다. 따라서 (1)에서와 같이 콘크리트 치기작업의 경우, 101동에서 안전사고가 발생하여 생산성 저하가 나타난 것을 알 수 있으며 각 동마다 자재조달 지연, 장비조달지연, 작업자 기능 부족 등에 따라 도출된 생산성 수치 값을 참고할 수 있다. 여기서, 생산성 조건으로 표현되는 정성적 요인의 경우 생산성 저하를 유발한 사안의 종류나 발생 정도에 따라 영향도의 편차가 매우 심할 수 있으므로, 충분한 data set이 확보되기 이전에는 그 값에 의미를 부여하기보다는 클레임관리 등을 위한 참고자료로써 그 결과물을 활용하는 것이 더욱 바람직 할 것으로 사료된다.

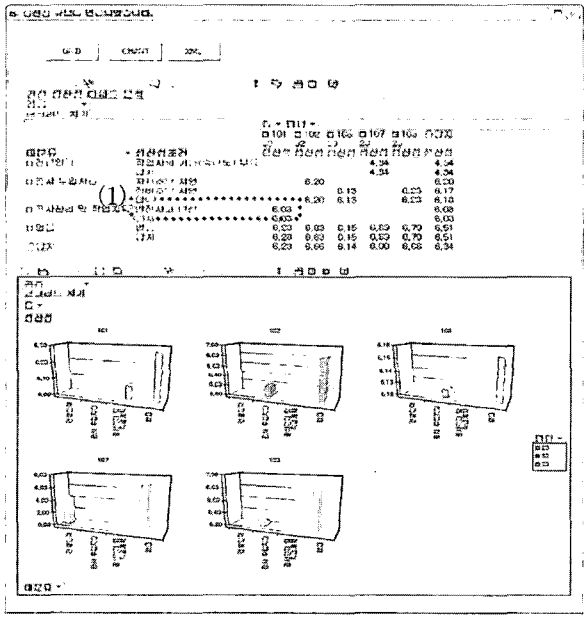

그림 13. 정성적 요인에 따른 생산성 분석

5.2 건설 생산성 예측 값의 제시

그림 14는 생산성 학습을 위한 화면으로 사용자는 공종을 선택하고 은닉층의 유니트 수 및 최대 학습회수를 입력한다. 또한 관리자의 판단에 따라 입력층에 반영할 생산성 영향요인을 선택할 수 있게 하였다. 이는 현장여건에 따라 생산성에 영향을 미치는 요인이 다를 수 있기 때문에 관리자는 OLAP의 다차원 분석을 통해 영향도를 판단하여 생산성 예측 모델에 반영할 요인을 선택함으로써 예측모델의 유연성을 제고하였다. 학습된 결과는 데이터베이스에 공종 별로 저장된다.

위와 같은 과정을 통해 특정 공종의 생산성에 대한 학습을 본 연구에서 제시한 인공 신경망 알고리즘에 적용하게 되면 그림 15와 같은 갱폼 인양에 대한 최적의 생산성 예측 값이 산정된다.

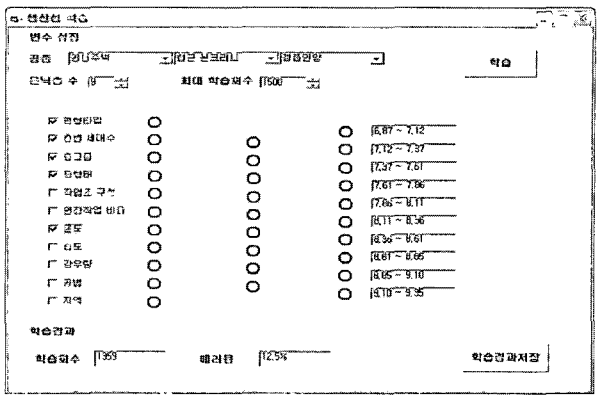

그림 14. 인공 신경망 알고리즘에 의한 생산성 학습

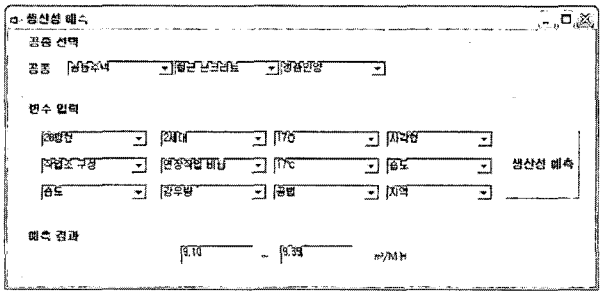

그림 15. 갱폼 인양의 생산성 예측 결과

그림 15와 같이 공동주택 철근콘크리트 공사의 갱폼 인양 작업에 있어 28평형, 2세대, 17층, 사각형태의 평면 구조와 온도가 17℃ 정도의 영향 조건을 설정할 경우, 약 9.10~9.35M2/M · H의 생산성 값이 산정됨을 알 수 있다.

이러한 생산성 값은 현재까지 축적된 실적자료를 토대로 본 연구를 통해 개발된 인공 신경망 알고리즘에 의해 산정되며, 여기에서의 실적자료는 도출된 영향요인을 포함하고 여러 현장에서 수집되어 데이터웨어하우스에 저장된 생산성 데이터를 의미한다. 단, 위 '9.10~9.35M2/M · H' 라는 값은 관리자가 공사계획에 반영해야 할 절대적 수치를 의미하는 것은 아니며, 유사 조건을 지니고 있더라도 현장 특성이나 관리자의 경험에 따라 수정보완(加 · 減)이 가능한 자사의 매우 객관적인 참고자료(reference data)임을 의미하는 것이다.

6. 결 론

본 연구에서는 건설 생산성 관리를 통한 체계적인 공사관리 및 합리적인 공사계획 수립을 위해 건설 생산성 영향요인을 도출하고 데이터웨어하우스, OLAP, 데이터 마이닝 기술 등을 활용하여 생산성 데이터를 축적, 분석 및 예측할 수 있는 건설 생

산성 관리 시스템을 제안하였다. 개발 시스템이 건설 현장에 적극 활용될 경우 건설 생산성 분석의 유연성 제고를 통한 건설 생산성 관리, 합리적인 건설 사업의 성과 측정뿐만 아니라 향후 유사 프로젝트에 있어 생산성 예측을 통한 보다 신뢰성 있는 공사 계획 등이 가능할 것으로 기대된다.

본 연구를 통해 얻은 결론은 다음과 같다.

1) 건설 생산성 데이터의 분석 및 예측을 위해, 본 연구에서는 건설 생산성 데이터를 노무 작업시간과 진행 물량의 비로 정의하였으며, 생산성 영향요인 분석과 관련된 문헌 고찰 및 현장 인터뷰를 통해 건설 생산성에 영향을 미치는 정성적, 정량적 요인을 제시하였다.
2) 기존의 RDB를 기반으로 하는 데이터베이스 구축체계를 데이터웨어하우스 체계로 전환하고 수집된 생산성 데이터를 ETL 과정을 거쳐 데이터웨어하우스에 축적하기 위한 방법론을 제시하였다. 또한 데이터웨어하우스에 축적된 생산성 데이터를 다차원적으로 분석하여 도출된 생산성 값과 영향요인간의 관계를 보다 유연하게 분석할 수 있는 OLAP 클라이언트를 구축하였다.
3) 신경망 알고리즘을 이용하여 생산성 영향요인에 따른 건설 생산성 예측 모델을 제안하였다. 이는 11개의 영향요인을 입력 층에 두고 10개의 생산성 수치를 출력 층으로 두어 가상으로 설정된 은닉 층을 통해 상호 관련된 값의 연결을 학습할 수 있게 하였으며, 이를 통해 건설 생산성 영향요인이 반영된 건설 생산성 예측 모델을 제안하였다.
4) 공동주택 구체공사를 진행 중인 표 4의 사례 현장으로부터 수집된 생산성 데이터를 적용하여 단위작업별 생산성 추이 분석, 협력업체별 성능 분석 등 시스템이 제공하는 주요 화면을 중심으로 도출된 생산성 값의 의미를 정의하였으며, 관계형 데이터베이스 및 OLAP 기반의 다차원적 생산성 분석의 결과를 비교·분석함으로써 OLAP을 기반으로 한 생산성 데이터 분석의 장점을 제시하였다. 또한 인공 신경망 알고리즘을 적용하여 관리자가 선택한 특정 작업환경 및 영향요인 하에서의 생산성 예측 값을 제시함으로써 건설 생산성 분석 및 예측을 위해 제안된 개발 시스템의 효용성을 검증하였다.

향후, 연구과제로서 다양한 공종 및 현장에 있어 개발 시스템을 적용하기 위한 수정·보완 작업이 요구되며, 이는 객관적이고도 신뢰성 있는 건설 생산성 실적자료의 구축을 가능케 하는 프레임워크를 제시할 수 있을 것이다.

감사의 글

본 연구는 2005년부터 진행된 과학재단 목적기초 연구과제 (R01-2001-000-00449-0)로 수행 되었으며 연구 지원에 진심으로 감사드립니다.

참고문헌

1. 김대수, "신경망 이론과 응용", 하이테크정보, 1992.
2. 김명호, "데이터웨어하우스를 이용한 건설 생산성 관리 시스템의 개발", 인하대학교 공학석사 학위 논문, 2005.08
3. 김예상, "건설 생산성에 영향을 미치는 요인 분석에 관한 연구", 대한건축학회 논문집, 제10권 10호, pp.267~273, 1994.10
4. 손창백, 이덕찬, "건축공사의 생산성 저하요인 분석", 대한건축학회 논문집, 18권 12호, pp.125~132, 2002.12
5. 안용선, "작업분석에 의한 철근 콘크리트조 거푸집 공사의 관리방안에 관한 연구", 한양대학교 박사학위 논문, 1993.02
6. 오세욱, 김영석, 이준복, 김한수 "PDA 및 바코드 기술을 이용한 건설 노무정보 수집 및 활용", 한국건설관리학회 논문집, 제5권 5호, pp.65~75, 2004.10
7. 오세욱, "단위작업 정보 기반의 공동주택 공정관리 지원시스템 개발", 인하대학교 공학박사 학위논문, 2005.02
8. 오세욱, 김영석, "단위작업 정보 기반의 공동주택 공정관리 지원시스템 구축방안에 관한 연구; 현장적용 및 검증결과를 중심으로", 한국건설관리학회 논문집, 제6권 6호, 2005.12
9. 유정호, "건설 프로젝트의 생산성 관리 시스템", 대한건축학회논문집, 제18권 7호, pp. 103~113, 2002.7
10. 조재희, "OLAP서버를 이용한 기업정보기반 구축에 관한 연구", 인문사회과학연구소 논문집, 25호, pp.275~291, 1996.
11. 표영민, "AHP기법을 이용한 건설 노동생산성 저하요인 분석에 관한 연구", 한국건축시공학회 학술 발표 논문집, 제8호, pp.141~149, 2005.05
12. BRE, "Calibre 2000", CD Rom for presentation 2000.
13. Han, J., and Kamber, M.,"Data Mining Concepts and Techniquew", Morgan Kaufmann Pubishers, 2001.
14. Inmon, W.H., "Building the Data Warehouse", Wiley, 1992.
15. Sonmez, R., and Napolitan, C.L., "Construction Labor Productivity Modeling with Neural Networks" Journal

of Construction Engineering and Management, Vol. 124,
No. 6, pp.399~410, 1995.
16. www.hostechglobal.com

논문제출일: 2005.08.12
심사완료일: 2006.03.15

Abstract

Productivity is important to evaluate an efficiency of performed work and organization in construction industry. The productivity should be defined as activity level rather than macro level in order to effectively use productivity data and manage a project. The primary objective of this study is to develop a construction productivity management system using data warehouse, OLAP and data mining technologies which enables to easily accumulate the construction productivity data and perform multi layer analysis. Finally, it is anticipated that the effective use of the developed system would be able to measure the result of project and make a plan of the similar project with reliability.

Keywords : Productivity, Data warehouse, OLAP, Data mining, influence factor