

공동주택 프로젝트의 초기 공사비 예측정확도 향상에 관한 연구

Improving the Accuracy of Early Stage Cost Estimation in Apartment Construction Project

임 소 연* 여 상 구** 고 성 석***
Lim, Soyeon Yeo, Sanggu Go, Seong-Seok

Abstract

Due to the diversification and complication of construction projects, controlling risks from the early design-planning phase gives huge impact on success of the construction project. As a part of managing uncertainties it is also important to estimate the project cost several times. Especially, estimating project cost in the early stage gives effects on making a budget for projects.

This study estimated the apartment project cost using case-based reasoning(CBR), which is the process of solving new problems based on the past problems. For this, we deduced the apartment cost influence factors which can be gathered in the early stage of project. Based on the factors we established the database for apartment project and calculated the attribute value, attribute similarity and case similarity. Although we retrieve the most similar case from the database, it is very hard to utilize it directly due to the uniqueness of each project. So, Genetic Algorithm(GA) was applied in revising the cost of the retrieved-case. Therefore, the accuracy of the prediction was improved by GA optimization.

키 워 드 : 공동주택 공사비, 사례기반추론, 유전자 알고리즘
Keywords : Apartment Cost, Case-based Reasoning, Genetic Algorithm

1. 서 론

건설 프로젝트는 초기·기획단계에서 소요예산이 측정되면 유효범위 내에서 효율적이고 실용적인 공법 및 대안을 결정함에 따라 전체 공사비용의 80%가 설계단계에서 결정되어져 설계 완료 후 조정 가능한 비용은 20%에 불과하다(P.Duverlie, 1995). 국내의 경우 공사비 정보를 사전에 파악하기보다 단순히 사후처리 정보로만 취급하여 공사비를 낭비하는 결과를 초래하고 있는 실정이며, 건설 분야에서는 동일한 용도의 건축물을 축조한다고 할 지라도 그 특수성으로 인하여 과거의 동일한 사례를 직접적으로 적용하기 어렵기 때문에 프로젝트 전 단계에서 지속적인 공사비 관리가 요구된다.

초기·기획단계에서의 공사비 예측이 효율적인 프로젝트 운영에 필수적임에 따라 회귀분석, 인공신경망, 유전자 알고리즘 등을 이용한 연구가 다양하게 진행되어 왔으며, 기수행되었던 공사의 데이터 분석이 요구되는 건설분야 특성상 경험 지향적 문제해결

기법을 이용하는 사례기반 추론이 공사비 예측에 가장 적합한 것으로 밝혀졌다(김광희 외, 2004). 반면 사례기반추론의 경우 새로운 문제와 가장 유사한 속성을 지닌 사례를 조회한 후 수정하는 방식을 적용하여 문제를 해결하기 때문에 공사비 예측에 효과적이지만, 기존의 데이터가 충분히 구축되어 있지 않거나 새로운 사례에 대한 문제해결을 해야 할 경우 적절치 못한 사례가 조회될 가능성이 있으며 속성가중치 산정 및 수정방안의 신뢰도가 결여되는 경우에는 예측정확도가 낮아지는 문제점이 발생하는 단점을 가지고 있다(김윤식, 2010). 특히 선행연구를 고찰한 결과, 선정된 해의 공사비를 새로운 문제의 속성에 근거하여 수정하는 방안에 대한 연구가 부족한 것으로 나타났다.

본 연구에서는 사례기반추론의 단점을 보완하여 공동주택 초기 공사비를 예측 정확도를 향상시키고자 사례기반추론에 의해 선정된 유사사례를 수정하는 보정단계에서 유전자 알고리즘을 적용하여 새로운 사례에 적합한 공사비로 수정하는 방안을 제시하였다.

* 전남대학교 건축공학과 석사과정, 정회원
** 전남대학교 건축공학과 박사과정, 정회원
*** 전남대학교 건축공학과 교수, 정회원

2. 이론적 고찰

2.1 사례기반추론

사례기반추론은 과거에 해결된 사례를 근거로 하여 새로운 문제의 해를 구하는 기법으로 유사한 문제는 유사한 해법을 갖는 다는 것과 한번 발생한 문제는 자주 발생된다는 가정에 기초하고 있다. 인간의 문제해결방식과 동일한 방식을 활용하여 도출된 결과를 이해하기 쉽고, 새로운 사례를 데이터베이스에 저장하는 것 이외에 추가적인 작업 없이도 학습을 진행할 수 있다는 장점을 지니고 있다(이재식, 2005). 사례기반추론은 Retrieve, Reuse, Revise 및 Retain 과정으로 해를 도출하며, 새로운 문제를 해결하는데 가장 적합한 사례를 도출해 내는 것에 따라 시스템의 정확도가 크게 좌우 된다. 그림 1은 사례기반추론의 문제해결 방법을 도식화 한 것이다.

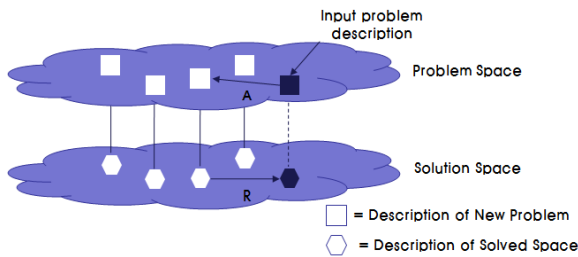


그림 1. 사례기반추론 문제해결 방식

1) Retrieve

새로운 사례와 유사한 특성을 갖는 과거사례를 데이터베이스에서 조회하는 과정으로, 유사사례를 조회하는 방법에는 귀납적 조회방법, 지식기반 조회방법, 최근이웃 조회방법이 있다. 귀납적 조회방법의 경우 목표가 명확히 정리되고 사례 데이터 자체에서 귀납된 결과에 대한 영향요소에 의하여 색인화 되므로, 최근이웃 조회방법보다 사례조회 속도가 빠른 장점을 가진다. 반면 사례데이터가 누락되면 사례조회가 전혀 불가능하게 되고, 사례가 증가할 때마다 의사결정 맵에 사례를 추가해야 되는 단점을 가지고 있다. 지식기반 조회방법은 규칙기반 시스템과 유사한 방식으로 사례를 조회하는 과정에서 기존의 분야별 지식을 적용한다. 마지막으로 최근이웃 조회방법은 데이터베이스에 축적되어 있는 사례들 중에서 유사성 척도에 의하여 새로운 사례와 유사한 사례를 조회하는 방법이며 특정 문제를 해결하는데 집중할 필요가 없는 경우에 적절한 방식으로, 사례의 수가 증가할수록 적합한 사례를 제시하는데 소요되어지는 시간이 증가하게 되므로 사례의 수가 다소 적은 경우에 적합하다.

2) Reuse

문제를 해결하기 위하여 조회된 사례를 이용하는 단계로 조

회된 사례를 재사용하기 위해서는 이 사례가 새로운 문제를 해결하는데 적합한지에 대한 충분한 고려가 필요하다.

3) Revise

가장 유사한 사례를 조회한 후 새로운 문제의 특성을 반영하기 위하여 기존의 해를 수정하는 과정이다. 본 연구의 범위인 공동주택 개발사업의 경우, 새로운 프로젝트와 가장 유사한 사례로 조회되었다고 할지라도 모든 속성이 동일하지는 않기 때문에 속성의 차이에 따라 공사비를 수정하는 방안에 관한 연구가 필요하다.

4) Retain

문제를 해결한 이후에 사례를 데이터베이스에 저장하여 차후 발생하는 문제에도 이용할 수 있게 하기 위한 과정이다. 생성된 사례의 성공 여부에 관계없이 관련된 정보는 이후에 사용되기 위하여 유효한 형태의 정보로 저장되어야 한다.

2.2 유전자 알고리즘

유전자 알고리즘은 자연진화의 법칙인 적자생존과 자연도태의 원리를 토대로 적용한 최적화 알고리즘이다. 유전자 알고리즘의 경우 적합도 평가, 재생산, 교배 및 돌연변이의 4단계를 거치게 된다. 적합도 평가는 개체 집단 내에서 적합도를 평가하고 그 결과에 따라 개체를 선택하는 과정이며, 이후 재생산에서는 적합도가 높게 평가된 개체를 선택하여 다음 세대에 전이한다. 두 부모 염색체의 교차점을 기준으로 서로 조합하고 새로운 개체를 만들어 낸 후 돌연변이에서 새로운 개체를 만들어 내게 되는데 진화가 거듭될수록 개체들은 적합도가 높은 개체들로 변화하게 되고, 최적해로 수렴하게 된다. 본 연구에서는 유사사례를 조회하여 획득된 공사비를 새로운 속성에 적합하게 수정하는 보정 단계에서 유전자 알고리즘을 적용하였다.

3. 영향요인 선정 및 데이터베이스 구축

2001년부터 2010년에 수행한 공동주택 37개 사례 중 3개를 제외한 34개를 수집하여 사례베이스로 구축하였다. 선정된 사례는 건설회사에서 자체 시행된 프로젝트이며, 실제 투입된 공사비의 직접비를 분석자료로 이용하여 공사비 예측의 정확도를 높이고자 하였다. 실적데이터를 공사비 예측 프로세스에 이용함으로써 사례기반추론에 의해 획득된 공사비 예측자료와 실제 투입원가를 비교하여 오차율의 여부를 판단할 수 있으며, 지속적으로 사례데이터를 구축할 경우 피드백이 가능할 것으로 사료된다. 37개의 프로젝트를 공사기간 완료시점과 공시지역으로 분류하여 나타내면 다음 표 1과 같다.

표 1. 사례데이터의 공사개요

	수도권	경북권	전라권	충청권	경상권	합계
2010	1	-	4			5
2009	4	1	4	1	-	10
2008	2			1	1	4
2007	4	-	3			7
2006	1	-	3		2	6
2005	1	-	-			1
2004	1	-	-			1
합계	14	1	14	2	3	34

공동주택 공사비 예측에 관한 선행연구를 고찰하여 총 37가지의 공사비 영향요인을 도출하였고, 견적전문가 및 개발사업 담당자와 면담조사를 실시하여 프로젝트 초기·기획단계에서 수집할 수 있는 공사비 영향요인으로 명목적 변수 3가지와 수량적 변수 12가지를 선정하였다. 사례데이터베이스는 그림 2에 제시된 방법으로 구축하였다.

순번	INPUT															공사비 (천원)	
	PHC 파일	기초 형식	기 보 형 식	매 감 수 준	연 면 적	대 지 면 적	건 축 면 적	주 차 장 면 적	건 폐 물	용 적 률	평 균 층 수	지 하 층 수	동 수	총 세 대 수	평 균 평 수		층 고
A1	1	0	0	1	43736	8592	2164	11199	25	379	20	2	5	319	96	2.7	28,736,122,498
A2	0	1	0	1	150458	30360	6807	40346	22	381	23	2	12	1060	104	2.7	91,937,050,728
A3	1	0	1	2	33897	12264	2109	2796	17	200	21	2	2	260	120	2.2	20,923,202,840

그림 2. 데이터베이스 구축

공사비 데이터의 경우 공사수행시기가 상이함에 따라 공사비 예측의 오차를 줄이고자 건설공사비 지수를 이용하여 보정하였다. 한국건설기술연구원에서 발표하는 건설공사비 지수를 이용하였으며, 2010년 1월을 기준년도로 설정하였다. 입력변수 역시 영향요인별로 단위가 상이하여 유사사례를 조회하는 과정에서 영향요인의 특성을 충분히 반영하지 못하는 문제가 발생할 수 있어 표준화된 수치로 변환하였다.

4. 사례기반추론에 의한 공동주택 공사비 예측

4.1 유사사례 조회¹⁾

본 연구에서는 최근이웃 조회방법을 이용하여 유사사례를 조회하였으며, 최근이웃 조회방법을 이용하여 유사사례를 조회하는 과정과 방법을 그림 3에 나타내었다.

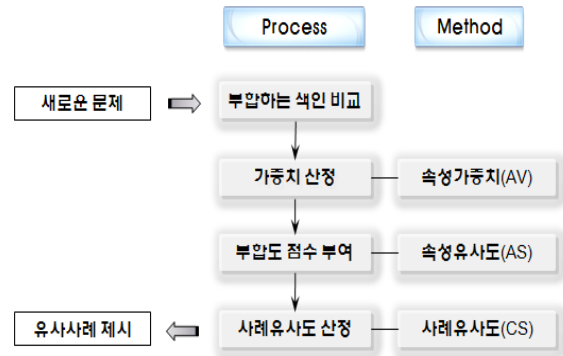


그림 3. 유사사례 조회과정¹⁾

1) 속성가중치 산정

속성가중치의 산정은 유사사례 조회에 적용하기 위하여 공사비에 미치는 영향을 토대로 가중치를 산출하는 단계이다. 유사사례의 예측정확도에 가장 밀접한 영향을 줄 수 있는 수치이기 때문에 선행연구자에 의하여 다수의 연구가 수행되었으며 본 연구에서는 회귀분석을 이용하였다. 앞서 3장에서 표준화된 자료를 이용하였으며, SPSS 12.0K를 이용하여 회귀분석을 실시하였다. 다중회귀분석의 경우 입력변수의 선택방법에 따라 입력선택, 전진선택, 후진제거, 단계선택이 있으며, 속성간 다중공선성의 문제를 해결하기 위하여 불필요한 변수를 삭제해가며 회귀식을 구성하는 후진제거를 적용하였다. 수정된 결정계수와 유의확률을 검토하여 최적의 모델을 도출하였고, 다음 표 3은 후진선택 방식에 따라 도출된 회귀식의 결과값과 수정된 R² 및 유의확률을 나타낸 것이다.

표 2. 후진선택 방식에 의해 도출된 회귀식

	회귀식	R ^{2adj}	유의확률
후진선택	$Y = 7,310,000,000 + 123,400,000,000x_1 + 1,990,000,000,000x_2 - 18,400,000,000x_3 - 19,000,000,000,000x_4 - 4,750,000,000x_5$ <p> x_1:연면적 x_2:대지면적 x_3:주차장면적 x_4:층수 x_5:평균평수 </p>	0.816	0.000

후진선택의 방법을 이용한 결과, 층고, 지붕형식, 평균평수, 지하층수, 동수, 연면적의 5개 변수가 영향요인으로 선정되었다. 수정된 R²값이 0.816으로 회귀식의 설명력이 충분하며, 유의확률이 0.000이기 때문에 통계적으로 유의함을 알 수 있다. 회귀분석의 결과값 중 표준화계수 베타값(β)의 절대값을 속성가중치로 이용하였다. 다음 표 3은 다중회귀분석을 통하여 산출한 각 영향요인별 속성가중치를 나타낸 것이다.

1) 이지은, "사례 기반 시스템의 이론적 고찰 및 구현현황", 경영학 연구 2 제6호, 상명대학교 경영연구소, 그림 재인용

표 3. 후진선택 방식에 의한 속성가중치

분류	속성가중치				
	연면적	대지면적	주차장면적	층수	평균평수
후진 선택	1,073	0.22	0.192	0.134	0.037

2) 속성유사도 및 사례유사도 산정법

새롭게 제시된 사례에 가장 적합한 유사사례를 조회하기 위해서는 사례유사도(Case Similarity, CS)를 산정하여야 하며, 이를 산정하기 위해서는 먼저 영향요인별로 속성유사도(Attribute Similarity, AS)를 산정하여야 한다. 속성유사도는 새로운 사례와 조회된 사례의 속성 즉, 영향요인간의 관계성을 알아보기 위한 것이다.

본 연구에서는 Euclidean의 거리 측정 개념을 이용하여 속성유사도를 산출하였다. 속성유사도와 속성가중치를 산정한 이후에는 이 값들을 조합하여 사례유사도를 산정한다. 각 영향요인별로 산정된 속성유사도(AV)와 속성가중치(ω)를 곱하여 합산한 값을 속성가중치를 합산한 값으로 나누며, 속성유사도를 산정할 시에 사례유사도(CS)를 산정한다. 사례유사도 산정방식은 식 (1)과 같다.

$$CS = \frac{\sum_{i=1}^n (f(N_i, S_i) \times AW_i)}{\sum_{i=1}^n \omega_i} = 1 - \sqrt{\frac{\sum_{i=1}^n \omega_i^2 (AV_{NCi} - AV_{RCi})^2}{\sum_{i=1}^n \omega_i}} \quad \text{식(1)}$$

3) 유사사례 조회기준

사례유사도가 가장 높은 사례가 새로운 문제와 가장 유사한 사업특성을 갖는다고 할 수 있기 때문에, 사례유사도를 기준으로 사례데이터에 저장되어 있는 사례를 정리한 이후에 사례유사도 점수 상위 1개의 사례를 조회한다. 그리고 해당하는 1개 사례를 1차적인 해결안으로 사용한다.

4.2 유전자 알고리즘을 이용한 유사사례 공사비 수정

건설 프로젝트마다 공사 속성값이 상이한 점을 고려할 때, 사례기반추론에서 조회된 사례와 새로운 사례의 속성 차이를 고려하여 조회된 공사비를 수정하는 과정이 필요하다. 본 연구에서는 유전자 알고리즘을 이용하여 저장된 사례의 속성별 차이가 공사비에 미치는 영향을 선형방정식으로 산출하여 조회된 해를 수정하였으며, 선형방정식 산출 방법은 다음과 같다. 먼저 유사사례로 선정된 사례와 사례데이터베이스에 저장된 사례들의 속성 차이가 사례별 공사비 차이에 미치는 영향을 일차식으로 나타내는 목적함수를 구성한다. 모든 사례에 대하여 속성차이와 공사비차이를 만족시키는 가중치(β)를 구하는 것은 불가능하므로 공사비차이와 각 해가 도출하는 수치와의 절대거리의 합을 최소화 하도록 최적화한다. 속성별 차

이를 반영하는 가중치를 새로운 사례와 조회된 사례의 속성 차이와 곱하여 합산한 후 속성별 변화에 따른 공사비 추이를 조회된 공사비에 반영한다. 본 연구에서는 유전자알고리즘의 사용화 프로그램인 Evolver 5.5.1을 이용하였다. 유전자 알고리즘을 이용하여 조회된 사례의 공사비를 수정하는 프로세스를 정리하면 아래 그림 4와 같다.

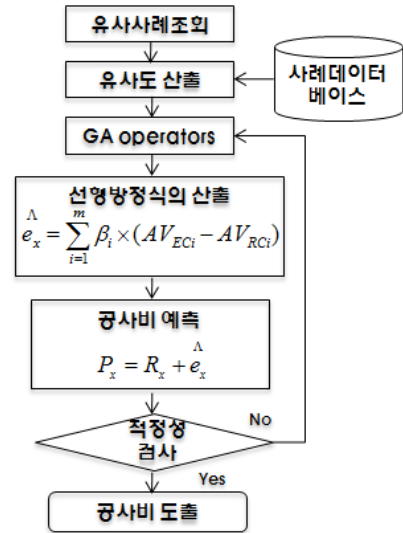


그림 4. 유전자 알고리즘을 이용한 공사비 수정 프로세스

4.3 검증 방법

본 연구에서 제시한 공동주택 공사비 예측방법을 검증하기 위하여 사례데이터베이스에 구축한 사례 중 무작위로 1가지를 선정한 후 선정된 사례 이외에 나머지 33가지의 사례는 데이터베이스에 구축되어 있는 상태로 총 5회에 걸친 검증을 실시하였다. 조회된 유사사례의 공사비와 유전자 알고리즘을 이용하여 공사비를 수정한 값의 검증결과를 표 4에 나타내었다. 사례 1을 대상으로 유사사례를 조회한 결과의 오차율은 18.94%로 나타났으나, 유전자 알고리즘을 이용하여 수정한 공사비의 오차율은 6.03%로 낮아졌다. 이외에도 검증 사례로 선정된 모든 사례에 있어서 오차율이 낮아진 것으로 보아 본 논문에서 제시한 방법의 유효성이 높은 것으로 나타났다.

표 4. 검증 결과

사례	해당공사비	CBR 예측결과		수정된 공사비	
		공사비	오차율	공사비	오차율
Case1	919천만	745천만	18.94%	863천만	6.03%
Case2	663천만	745천만	12.27%	607천만	8.47%
Case3	228천만	287천만	25.88%	201천만	11.90%
Case4	745천만	663천만	10.93%	811천만	8.91%
Case5	1410천만	919천만	34.80%	1110천만	21.22%
평균			20.56%		11.31%

5. 결 론

본 연구에서는 초기·기획단계에서의 공사비 예측이 효과적인 사업추진에 있어서 매우 중요한 요소로 작용할 수 있다는 점에 착안하여, 건설회사에서 자체 수행한 공동주택 프로젝트를 대상으로 초기·기획단계의 공사비를 예측하였다. 기수행되었던 공사의 실적공사비 데이터를 비교분석하여 새로운 프로젝트의 공사비에 가장 유사한 사례를 조회하는 사례기반추론을 이용하였으며, 조회된 사례의 공사비를 해당 사례의 속성에 맞게 수정하는 단계에서는 유전자 알고리즘을 이용하여 보정하였다. 본 연구에서 도출된 결론은 다음과 같다.

선행연구 고찰을 통해 공사비 조회한 영향요인 중 전문가 면담을 통해 명목적 변수 3가지와 수량적 변수 12가지를 선정하여 데이터베이스를 구축하였다. 유사사례조회 방식을 이용하여 유사사례를 조회하였으며, 그 과정에서 회귀분석의 후진 제거 방식을 이용하여 5가지의 영향요인을 최종적으로 이용하여 속성 가중치를 산출하였다. 회귀분석의 결과를 살펴보면 연면적이 가장 높은 가중치를 나타냈으나, 다른 영향요인도 영향요인으로 적용하게 되므로 연면적을 단일로 이용하여 공사비를 산출하는 방식보다 공사의 특성을 반영할 수 있을 것으로 사료된다. 조회된 유사사례의 공사비를 유전자 알고리즘을 통해 산정된 1차식을 토대로 수정한 결과, 수정된 공사비의 경우 사례기반추론을 통해 조회된 유사사례보다도 평균 오차율이 20.56%에서 11.31%로 감소함에 따라 보다 정확도가 높은 공사비를 예측할 수 있을 것으로 기대된다.

본 연구의 한계점으로는 사례가 충분하지 못하여 유사사례 조회 및 유전자 알고리즘을 통해 산출된 일차식의 검증에 다소 한계가 있었으며, 지속적인 사례 구축을 통한 보완이 필요할 것으로 사료된다.

감사의 글

이 논문은 2010년 교육과학기술부로부터 지원받아 수행된 연구임(지역거점연구단육성사업/바이오하우징연구사업단)

이 논문은 2010년 전남대학교 건축과학기술연구소 및 바이오하우징 연구소의 지원을 받아 수행된 연구임

참 고 문 헌

1. 구충완, 공동주택 사업특성에 따른 공사기간 및 비용예측을 위한 CBR 기반 Hybrid 모델, 서울시립대학교 대학원 석사학위논문, 2007
2. 김광희 외, 공동주택 공사비 예측 정확도 비교에 관한 연구, 대한건축학회논문집, 제20권 제5호, 2004.5
3. 김광희 외, 공동주택 프로젝트의 초기공사비 예측을 위한 신경망 학습

- 에 유전자 알고리즘을 적용한 모델에 관한 연구, 대한건축학회논문집, 제19권 제10호, 2003.10
4. 김광희 외, 사례기반추론 기법을 이용한 공동주택 초기 공사비 예측에 관한 연구, 대한건축학회논문집, 제20권 제5호, 2004.5
5. 김문한 저, 건설프로젝트의 코스트매니지먼트, 국토일보, 2009
6. 김수영 외, 사례기반추론의 정성변수 속성가중치 선정방법: 건축공사 초기 공사비 예측을 중심으로, 한국건설관리학회 정기학술발표대회 논문집, 제9권, 2009
7. 김윤식 외, 사업초기단계의 공사비 예측을 위한 CBR기반 MRA 보정 모델, 서울시립대학교 대학원 석사학위논문, 2010.
8. 지창윤, 공동주택 건설공사의 공사비 예측정확도 향상을 위한 CBR-Revision Model, 서울시립대학교 대학원 석사학위논문, 2009
9. A. Aamodt, E. Plaza, "Case-Based Reasoning Foundational Issues, Methodological Variations, and System Approaches.", AI Communications, IOS Press, Vol,7, No,1, 39~59, 1994.
10. Duverlie and J. M. Castelain, "Cost Estimation During Design Step: Parametric Method versus Case Based Reasoning Method", Int J Adv Manuf Technol, 15 895~906, 1999