Institute of Internet and Intelligent Technologies
Vilnius Gediminas Technical University
Saulėtekio al. 11, 10223 Vilnius, Lithuania
http://www.isarc2008.vgtu.lt/

The 25th International Symposium on Automation and Robotics in Construction

June 26–29, 2008

ISARC-2008

COST ESTIMATION METHODOLOGY USING DATABASE LAYER IN CONSTRUCTION PROJECTS

Kyeong Jin Han
Dept. of Architecture
Seoul National University
39- 425 Seoul National University, Sillim-dong
Kwanak-gu Seoul, Korea, 151-742
hkjtop@snu.ac.kr

Hyun-Soo Lee
Dept. of Architecture
Seoul National University
39-430 Seoul National University, Sillim-dong
Kwanak-gu Seoul, Korea, 151-742
hyunslee@snu.ac.kr

Moonseo Park
Dept. of Architecture
Seoul National University
39-433 Seoul National University, Sillim-dong
Kwanak-gu Seoul, Korea, 151-742
mspark@snu.ac.kr

Sae-Hyun Ji
Dept. of Architecture
Seoul National University
39-425 Seoul National University, Sillim-dong
Kwanak-gu Seoul, Korea, 151-742
oldclock@snu.ac.kr

ABSTRACT

Recently, as the scale, structure, capacity and use of buildings are more complicated in construction projects, it is more difficult to estimation project's costs. Because its project's scale is much bigger than any other project's one and total project cost is concluded after project is finished, it is important to make cost estimation method and system. In past, it is possible to estimate cost only using simple method. But, to cope with the now and future situation, it is necessary to have to study not only method but also database. In this paper, using real structure data of Korea public residential buildings, we discuss cost estimation based on the quantity of material at design stage. To estimate costs accurately under complicated situations, After selecting influence factors, cost database, named Database Layer, is created. And then, cost model is suggested using Case-Based Reasoning and Regression. This paper is expected to improve the accuracy and contribute to estimation studies under various conditions.

KEYWORDS

Cost estimation, Cost database, Influence factor, CBR, Regression

1. INTRODUCTION

Construction projects have characteristics that one project's scale is much bigger than any other project's one and total project cost is concluded after project is finished. Moreover, even though projects' scale and use are same, their total cost is changed by region, site, owner's requirement, culture, economic condition and so on. That is, at design stage, a perfect cost is not able to be concluded. But, it is important to make a conclusion for a project decision making. Recently, as the scale, structure, capacity and use of buildings are more emphasized, cost estimation is more and more difficult. Therefore, it is helpful to estimate an accurate cost and decrease a period of the cost estimation for the organization's decision making and successful cost management. The general cost estimating method is to select and modify similar cases based on past data. Because of the characteristic of this method, if the number of data is not sufficient, it is not easy to estimate the cost accurately. Therefore, desirable cost estimation methodology should even satisfy not good conditions. In order to develop the methodology on design stage, this paper suggests the cost estimation methodology using CBR (Case-based reasoning), Regression and the concept of Database Layer.

2. LITERATURE REVIEW

2.1. Terminologies of Cost Management

Cost management is processes involved in cost planning, estimating, budgeting and controlling so that the project can be completed within the approved budget for succeeding a project [1]. In its process, cost management is divided into cost planning, cost control and cost planning is, one more, divided into cost estimating, cost budgeting and cost check (Figure 1).

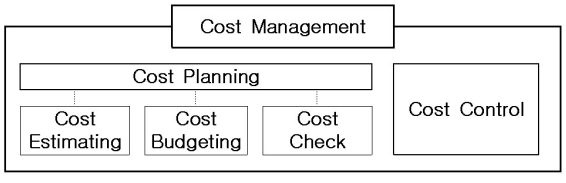

Figure 1. Cost management elements

Cost planning aims to help all members of the design team to arrive jointly at practical and efficient designs for the project and to keep within the budget [2]. Cost estimating is to developing an approximation of costs of the resources needed to complete project activities and cost budgeting is to aggregating the estimated costs of individual activities or work packages to establish a cost baseline [1]. Cost check, as definition of preconstruction for cost control, is the total process which ensures that the contract sum is within the client's approved budget or cost limit. Last, cost control aim at ensuring that resources are used to best advantage [2].

2.2 Process of Cost Management

The Australian Institute of Quantity Surveyors defines cost management process to 6 stages in Australian Cost Management Manual (Figure 2) [3]. In AIQS process, B stage includes feasibility study and C does schematic design in general. Construction stage generates fixed data, which are used to cost estimating in cost planning.

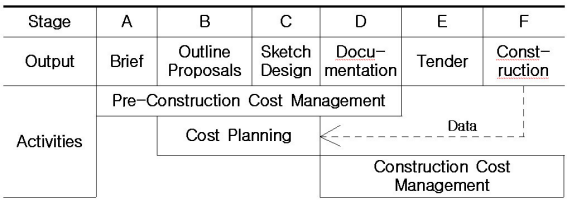

Figure 2. Cost management process

2.3 Degree of Accuracy for Cost Estimating

In early pre-construction stage, accuracy degree is relatively low because of insufficient information. But, as project is going on and amounts of information are more gained, it is generally improved. Some organizations have degree acceptable for each stage (Table 1) [4]. But these degrees may be changed by characteristic, period and area of project.

Table 1. Degree acceptable for cost estimating

AACE		ACostE		PMI	
Stage	Accuracy (%)	Stage	Accuracy (%)	Stage	Accuracy (%)
Class 5 (Order of magnitude)	±25 ~ 100	Order of magnitude	-30 ~ +30	Class 4	+50 ~ -30
Class 4 (Study)	±15 ~ 50	Study	-20 ~ +20	Class 3	+30 ~ -15
Class 3 (Preliminary)	±10 ~ 30			Class 2	+15 ~ -10
		Budget	-10 ~ +10		
Class 2 (Definitive)	±5 ~ 15	Definitive	-5 ~ +5	Class 1	±5
Class 1 (Detailed)	±3 ~ 5				

2.4. Case-Based Reasoning for Cost Estimating

Although each construction project has a unique characteristic, it also has the similarity. Therefore, Case-Based Reasoning (CBR) is suitable for analysis of construction cost, especially residential building. CBR is a method to solve a new problem by remembering a previous similar situation and by reusing information and knowledge of that situation [5]. CBR has a goal of enhancing computer intelligence and a characteristic more human-like than other methods. A general CBR cycle is described by four processes: 1) RETRIEVE the most similar case or cases 2) REUSE the information and knowledge in that case to solve the problem 3) REVISE the proposed solution 4) RETAIN the parts of this experience likely to be useful for future problem solving (Aamodt, 1994) (Figure 3).

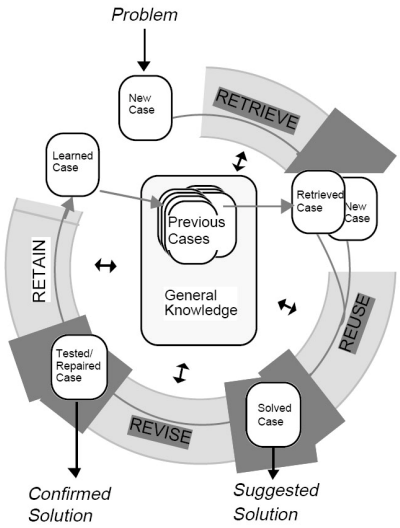

Figure 3. The CBR Cycle (Aamodt, 1994)

3. DATABASE ARCHITECTURE FOR COST ESTIMATION

3.1. Acquisition of Previous Data

In order to estimate new project cost, it is essential to acquire previous data. This paper uses structure quantity data of Korea public residential project which have common and similar features (Table 2)

Table 2. Basic features of previous projects

Figure	Value
Unit size	49, 59, 84, 114 m2
Floor	11~15
Underground floor	-1
Structure	RC
Floor height	2.8~3m

3.2 Selection of Items to Use

Bill of quantities (BOQ), regarding Korea structure cost data, consist of 3 major items, concrete, steel, and form. The total cost of these items in structure account for about 99.4% (Table 3). Therefore, to estimate cost efficiently, only three major items is estimated.

Table 3. Cost ratio of items in structure

Items	Cost Ratio
Concrete	24.51%
Steel	25.17%
Form	49.73%
Others	0.59%
Sum	100%

3.3. Influence Factors

To retrieve previous data from database, it is very important to establish influence factors which have detail features for data. They are used to the input process of a new case (data) and the part of database architecture. They are concluded by interacting information including basic features, drawings. Since Korean residential buildings are similar patterns, the smallest analysis unit is not house units but buildings. After analyzing information variously, 5 factors were selected, main unit size, the feature of the building plan, the number of units per a floor, the number of floor and void that reveal open area ratio on the drawing elevation (Table 4).

Table 4. Influence factors

Influence factors	Partitions
Main unit size	49, 59, 84, 114 m2
The feature of the building plan	L-shape, bar-shape
Units_(core) per a floor	2_(1), 4_(1), 4_(2)
The number of floor	11~15
Void (open units / total units)	0~100%

3.4. Database Architecture

Based on selected items and influence factors, the input fields of previous cases, or database, consist of as shown in table 5. Viewed at function's angle, factor values are independent variables, and Item values are dependent ones. Previous cases named and stored by user have to be accumulated in database frame. Figure 4 shows method that can fill cases(data) with influence factors. First, each data is input into Layer by main unit size. Then, in Layer, it is placed on X-Y coordinate by the number of floors and unit_(core) per a floor. Last, the data placed by a upper method has void value by itself. Figure 5 show an example applied by this method.

Table 5. The input fields of previous cases and examples

Input influence factors	Value		
Name	K101	K102	…
Main unit size (m2)	59	84	…
The feature of the building plan	bar	L	…
Units_(core) per a floor	4_(2)	4_(1)	…
The number of floor	15	14	…
Void (open units / total units)	10.00 %	6.67 %	…
Input items	**Value**		
Concrete (m3)	3,430	5,210	…
Steel (ton)	328	517	…
Form (m2)	28,961	41,112	…

Layer (Main unit size)					
	Units_(core) per a floor				
The number of floors	Blank	Blank	Blank	Blank	Blank
	Blank	**Fill**	Blank	**Fill**	Blank
	Blank	Blank	Blank	Blank	Blank
	Blank	**Fill**	Blank	Blank	Blank
	Blank	Blank	Blank	Blank	Blank

Figure 4. The concept of DB architecture

84m2			
Type	Bar-Shape		L - Shape
	2_(1)	4_(2)	4_(1)
11		K208_00.00	K205_09.09
12	K201_00.00 K703_00.00 K1002_00.00	K409_08.33 K509_08.33 K805_08.33 K202_00.00 K203_00.00	K102_04.17 K803_08.33 K804_08.33
13			K403_03.85 K911_07.69 K1003_03.85 K1006_03.85 K1008_03.85
14		K507_03.57 K801_10.71	K710_07.14
15	K610_06.67 K711_06.67		K406_06.67 K601_10.00 K602_13.33 K603_13.33 K604_13.33 K708_06.67 K709_06.67 K910_06.67

Figure 5. The example of DB architecture

4. COST MODELING METHOD FOR COST ESTMATION

Once database was established, various methodologies can be applied to estimating as the database form. Analyzing the suggested form, there are two kinds of cells: one is Fill, the other is Blank. If a problem that has to be estimated has an accurate influence factors with previous cases, data for solution is retrieved directly and rightly. In the other hand, if not, it may be reasoned by around data using Regression.

4.1. Cost Modeling for Same Database Layer

When problems are estimated for budget, if previous data are filled perfectly, it will be very easy to decide future budgets for projects. However, in reality, we sometimes, or always, face up to situations to estimate cost with insufficient data. Therefore, the ideal cost model has to satisfy two conditions: one is that it must have a higher accuracy with given data, the other is that it has to explain how the results were estimated [6]-[9]. Figure 6 shows how a problem is estimated in cost model. There are some filled data in database in it. (a) shows that this case can acquire a previous case directly because of the accurate matches of influence factors. In the other hand, (b) and (c) must use regression method. Naturally, accuracy degree of (a) is the highest of the others. But through (b) and (c), those who estimate the cost can grip foundations for the results. Simultaneously, since this method uses all given data, it can have a high accuracy.

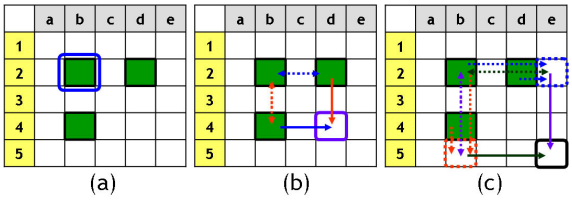

Figure 6. The concept for cost model of same DB layer

4.2. Cost Modeling for Different Database Layer Application

In this paper, database layers are divided by 'main unit size'. Like cases of same database layer, different layer also have a possibility to show the imbalance of data. The method for cost estimation is almost same with same layer's one (Figure 7). Going one step forward, if a standardization of this database frame is fulfilled, it is also possible to extend and create database layer.

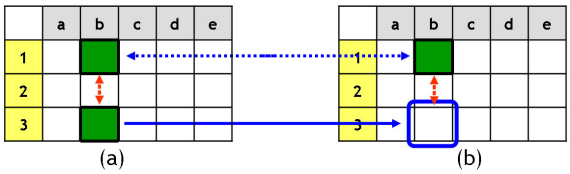

Figure 7. The concept for cost model of different DB layer

5. VERIFICATION

In order to verify the suggested method, the data in figure 5 were used. The verification method is to compare the real data of 'K610_06.67%(2_1, 15)' with its estimated value that use 'K201(2_1, 12)', 'K102(4_2, 12)' and 'K406(4_2, 15)'. The procedure is followed by three processes: First, void values of all data to use are changed to zero. Next, the concept of figure 6(b) is used. Last, the result is compared. As a result of this, Accuracy rate was calculated to +5.31% (Table 6(a)), and Applying cost ratio, Accuracy is improved to +3.64% (Table 6(b)).

Table 6 (a). Verification result

Items	Real Data	Estimated Value	Accuracy
Concrete (m3)	2437	2292	+6.36%
Steel (ton)	245	226	+8.54%
Form (m2)	19518	19313	+1.06%
Average			+5.31%

Table 6 (b). Verification result

Items	Accuracy	Cost ratio	Cost Accuracy
Concrete (m3)	+6.36%	0.245	
Steel (ton)	+8.54%	0.252	
Form (m2)	+1.06%	0.497	+3.64

In conclusion, comparing this result, +3.64%, with AACE's accuracy degree (Table 1), it is convinced to reach a rational result [10].

6. CONCLUSIONS

This paper discussed the cost estimation at design stage. Especially, after gathering real data, database layer was created by selecting and analyzing influence factors and then, based on this database, cost model to use CBR and Regression was suggested. And through verification, we could acquire a remarkable result.

In fact, if amount of data is sufficient, we can estimate projects' costs easily using only basic CBR. Besides, we may not need well-made database, either. But, a reality is different. Therefore, it is important to study and make a system to be able to use any situation.

In the future, through more study, it will be necessary to fix the theory and to make an automated system for users.

ACKNOWLEDGEMENT

This research was supported by a grant (R&D06CIT-A03) from Innovative Construction Cost Engineering Research Center, and (05CIT-D05-01) from Construction Technology Innovation Program funded by Ministry of Land, Transport and Maritime Affairs of Korean government.

REFERENCES

[1] PMI Project Management Body of Knowledge 3rd edition, PMI, 2004.

[2] Seeley, I.H., (1997). Quantity Surveying Practice 2nd edition, Antony Rowe Ltd.

[3] AIQS Australian Cost Management Manual Volume 1 3rd edition, AIQS, 2006.

[4] Ahn, Y. S., (2003). Improving the Accuracy of Screening of Cost Estimating in Early Construction Project Phase, Korea Journal of Architectural Institute, Vol.19, No.11, 2003, p. 133–140.

[5] Aamodt, A.; Plaza, E., (1994). Case-Based Reasoning: Foundational Issues, Methodological Variations and System Approaches, Al Communications, Vol. 7, No.1, 1994, p. 39–59.

[6] Kim, K. D., (1990). A Study on the Rational Cost Model of Apartment House in the Incipient Planning Stage, Korea Journal of Architectural Institute, Vol.6, No.3, 1990, p. 291–298.

[7] Liu, L.; Zhu, K., (2007). Improving Cost Estimates of Construction Projects Using Phased Cost Factors, Journal of Construction Engineering and Management, ASCE Vol.133 No.1, 2007, p. 91–95.

[8] Son, B.S., (2007). Development of Cost Estimating Model Using Quantity Variation Analysis by Influence Variables for the Schematic Design Phase, Korea Journal of Construction Engineering and Management, Vol.8, No.2, 2007.

[9] Yau, N. J.; Yang, J. B., (1998). Case-Based Reasoning in Construction Management, Computer-Aided Civil and Infrastructure Engineering, Vol.13, 1998, p. 143–150.

[10] AACE Cost engineers' notebook, AACE, 1995.