S17-3

COST ESTIMATE AT EARLY STAGE USING CASE-BASED REASONING

Kihoon Seong[1], Moonseo Park[2], Hyun-Soo Lee[3], and Sae-Hyun Ji[4]

[1] Master Course Student, Seoul National University, Seoul, Korea
[2] Professor, Seoul National University, Seoul, Korea
[3] Professor, Seoul National University, Seoul, Korea
[4] PhD Student, Seoul National University, Seoul, Korea
Correspond to emperorskh@hotmail.com

ABSTRACT: The importance of cost estimate in early stage such has been increasing due to market change and severe competition in construction industry. Because the adjustable budget is only 20% after design stage, most of the crucial decisions to influence cost is made in the early stage. However, in the early stage, the project scope is not defined completely so that estimator has inaccurate information to make critical decision. Therefore, this research suggests the cost estimate method using case-based reasoning. Case-based reasoning is appropriate for the early cost estimating, as it has the strength of rapidity and convenience in cost estimation. This research analyzes 84 actual data of public apartment on the scale of 11～15 stories. In order to extract the most similar case, at the first step this research identifies influence factors and calculates attribute similarity. In case-based reasoning, the most challenging task is determining attribute weight. At the third step, this research calculates case similarity which is aggregated attribute similarity multipled by attribute weight. Finally, extracts the most similar case which has the highest score of case similarity.

Keywords: Cost Estimate, Early Stage, Case-Based Reasoning, Nearest Neighbor Retrieval, Genetic Algorithm

1. INTRODUCTION

1.1 Background and Objective

Recently, market change and severe competition become more serious in the construction industry. Because of this situation, cost is considered to be critical success factor of construction projects. Particularly, cost at early stage such as conceptual and schematic stage has considerable degree of influence on total construction cost. At the early stage, about 80% of total construction cost can be changed. However, after design stage, the adjustable cost is only 20% of the total cost(P. Duverile, 1995). Moreover, accurate and trustworthy cost estimate is necessary for decision-making. However, because design information is not available at the conceptual and schematic stage, it is difficult for decision-maker to reach correct judgment. Cost estimate at early stage is usually performed based on unit cost per area(㎡). This kind of cost estimate has disadvantages of large amount of error and lack of connection with next stage such as detailed design. Reliable and accurate cost estimate model is essential for deciding appropriate budget.

Therefore, this research suggests the cost estimate model for early stage using case-based reasoning. The suggesting cost model is systematically organized and can reflect the increase of information. As the major technique in this cost model, case-based reasoning is one of Artificial Intelligence method, and is utilized in the various fields such as construction and diagnostics.

1.2 Scope and Methodology

Generally, pre-construction process is divided into feasibility study, conceptual design, schematic design, detailed design and procurement/contracting. Among these stages, this research focuses on conceptual design stage. This research dealt with the mid-rising public apartment projects in South Korea. The project which are analyzed and validated in this research is comprised of 9 complex, 84 units of apartment. And all units of apartments are on the scale of 11~15 stories.

In order to suggest the cost estimate model using case-based reasoning, this research applied the following procedure.

First, it examines the principle and application of case-based reasoning and genetic algorithm through studying previous researches.

Second step is to analyze data for developing cost model and verifying the reliability of it.

Third, it develops the cost estimate model using case-based reasoning and genetic algorithm as the major techniques.

Finally, this research examines the validity of this cost estimate model. Ultimately it confirms which this model can be utilized as the means of early cost estimate.

The figure1 shows the procedure of this research.

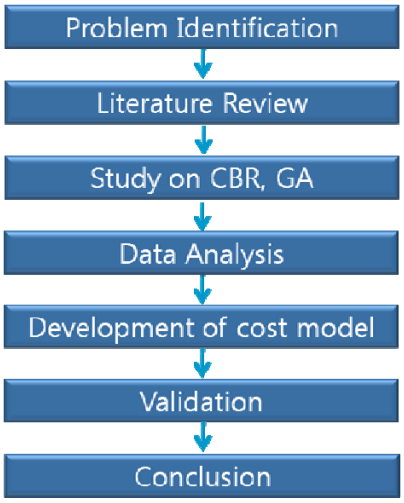

Fig. 1. Research procedure

2. BACKGROUND OF CASE STUDIES

2.1 CASE-BASED REASONING, CBR

2.1.1 Concept of Case-Based Reasoning

Case-based reasoning(CBR) is a problem solving approach that in many respects is fundamentally different from other major AI techniques such as neural networks and experts systems. CBR is able to utilize the specific knowledge of previously experienced problem situations. A new problem is solved by finding a similar past case, and reusing it in the new problem situations(Aamodt and Plaza 1994). CBR contains two main assumptions, which are that similar problems have similar solutions and once happened problem tends to come about again. In other words, CBR basically solves a new problem by remembering a previous similar situation and by reusing information and knowledge of past situation(Watson and Marir 1994). The CBR is utilized in the field of construction claim(Arditi et al 1999), cost estimate, diagnostics(Koton 1998), decision-making on bidding(Chau et al. 2001). Espeically for diagnostics, the application of past experience is important for solving the early stage problem(Lopez & Plaza 1997).

2.1.2 Case-Based Reasoning process

Case-based reasoning is composed of following 4 steps to suggest a solution for the present problem based on the past experience.

The first step is retrieving, which means that to extract similar previously experienced cases whose problem is judged to be similar. The second step is reusing the cases by copying or integrating the solutions from the cases retrieved. The third step is revising or adapting the solutions retrieved in an attempt to solve the new problem. The last one is retaining the new solution once it has been confirmed or validated(Aamodt and Plaza 1994).

The figure2 shows the case-based reasoning process.

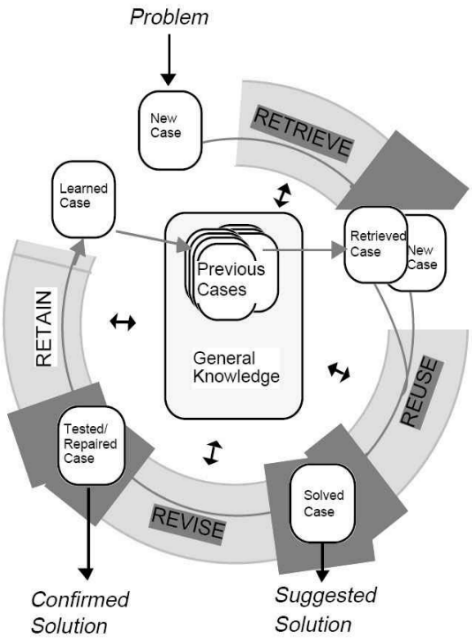

Fig. 2. CBR process

2.1.3 CASE RETRIEVAL METHOD

a. Nearest-neighbor retrieval

In nearest neighbor retrieval, the case retrieved is chosen when the weighted sum of its features that match the current case is greater than other case in the case base. Features that are considered more important in a problem-solving situation may have their importance denoted by weighting them more heavily in the case-matching process.

Especially, the K-nearest neighbor principle involves search for the K nearest cases to the current input case using a distance measure, and then selecting the class of the majority of these K cases as the retrieval one.

b. Inductive retrieval

Induction is a technique developed by machine learning researchers to extract rules or construct decision tree from past data. In case-based reasoning, the case-base is analyzed by an induction algorithm to produce a decision tree that classifies the cases. When inductive approaches are used to determine the case-base structure, which determines the relative importance of features for discriminating among similar cases, the resulting hierarchical structure of the case base provides a reduced search space for the case retrieval. This may, in turn, reduce the query search time.

c. Nearest neighbor retrieval vs. Inductive retrieval

Nearest neighbor is a simple technique that provides a measure of how similar a target case is to a source case. Nearest neighbor has one major weakness, namely, retrieval speed. Inductive retrieval also has one major disadvantage. If case data is missing or unknown, it may not be possible to retrieve a case at all. Nearest However, nearest neighbor retrieval is much less sensitive to missing, or noisy, case data. Even though there is not the most exact or suitable data, nearest neighbor retrieval

would still work and might be able to recommend the most similar case.

Table 5. comparision between two method

	Nearest Neighbor	Inductive method
Strength	less sensitive to noisy or missing case data	quick retrieval time
Weakness	retrieval speed	sensitive to missing or unknown case data

2.2 Genetic Algorithm, GA

2.2.1 Concept of Genetic Algorithm

Genetic algorithms are search algorithms based on the mechanics of natural selection and natural genetics. They combine survival of the fittest among string structures with a structured yet randomized information exchange to form a search algorithm with some of the innovative flair of human search. In every generation, a new set of artificial creatures(strings) is created using bits and pieces of the fittest of the old. While randomized, genetic algorithms are no simple random walk. They efficiently exploit historical information to speculate on new search points with expected improved performance.

2.2 2 Genetic Algorithm process

Genetic Algorithm is composed of three operators; reproduction, crossover, mutation.

Reproduction is a process in which individual strings are copied according to their objective function value. Copying strings according to their fitness values means that strings with a higher value have a higher probability of contributing one or more offspring in the next generation.

After reproduction, crossover my proceed in two steps. First, members of the newly reproduced strings in the mating pool are mated at random. Second, each pair of strings undergoes crossing over.

Mutation is the occasional random alteration of the value of a string position. Even though reproduction and crossover effectively search and recombine extant notions, mutation is needed because occasionally they may become overzealous and lose some potentially useful genetic material.

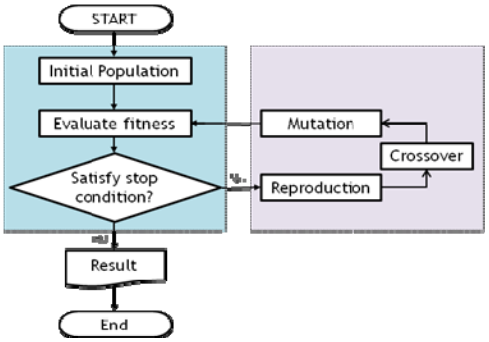

Fig. 3. Genetic Algorithm process

3. DATA ANALYSIS

3.1 Data Summary

The collected data is divided into 4 area types of 49 ㎡, 59 ㎡, 84 ㎡, 114 ㎡. As referred above, data is composed of 9 complex and 84 units of apartments. Because the number of 114 ㎡ area type is not enough to use, 114 ㎡ area type is excluded from this research. Table2 below shows the constitution of area type in each apartment complex.

Table 6. Data Summary

	n	n+1	n+2	n+3	n+4	n+5	n+6	n+7	n+8
49 ㎡			3	3			2	1	1
59 ㎡	4	1		1	2	2	2	2	3
84 ㎡	2	5	5	3	6	5	5	2	4
114 ㎡		1						5	3
49+59 ㎡			1	1					
49+84 ㎡			1					3	
59+84 ㎡	1			1					1
84+114 ㎡		2							
etc.					2	4			
Total	7	9	10	9	10	11	9	13	12

3.2 Data Analysis method

The chosen project shows various of shape, household combination and change of shape. Despite of the same area type of apartments, they have different shape and number of piloti household and so on. One unit of apartment has many of factors to influence on cost. Therefore, this research analyzed data by dividing unit of apartment based on area type and shape. It estimates individual cost for each area type instead of assess the total cost for whole apartment. For example, in the figure4 below, according to the data analysis method, this research classifies one unit of apartment into 84 ㎡ and 49 ㎡ area type. By using this analysis method, this research can reflect change of shape, area type and top floor.

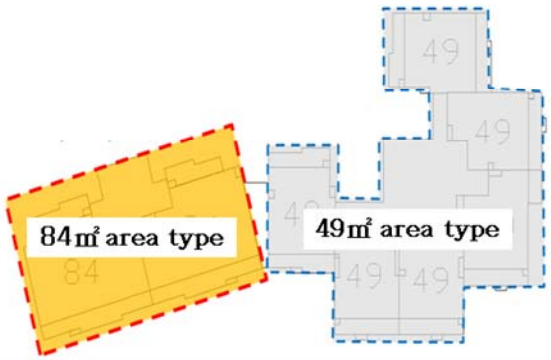

Fig. 4. Dividing area type in an apartment

Table 7. Established database

unit	area(㎡)	# of household	GFA	combination of household	Floor	# of elevator	# of household per elevator	# of piloti household	Total Cost
k101	59	46	5065.7	4	12	1	4	4	2,691,113,518
k101	59	22	2452.91	2	12	1	2	4	1,303,089,259
k102	84	46	5065.7	4	12	1	4	4	2,670,302,790
k103	59	20	1654.12	2	11	1	2	4	912,157,072
k103	84	42	4629.73	4	11	1	4	4	2,553,043,890
k104	59	38	3103.27	4	10	2	2	0	1,796,698,184
k105	59	20	1648.42	4	5	2	2	0	1,082,968,204
k106	59	30	2472.63	6	5	3	2	0	1,563,441,792
k107	59	18	1508.43	4	5	2	2	4	1,068,383,339
k201	84	24	2638.56	2	12	1	2	0	1,478,020,520
k202	84	48	5277.12	4	12	2	2	0	2,647,914,525
k203	84	48	5277.12	4	12	2	2	0	2,629,430,221
k204	59	46	3749.87	4	12	2	2	2	2,034,718,903
k205	84	40	4448.13	4	11	1	4	4	2,398,184,147
k206	84	48	5320.07	4	13	1	4	4	2,762,769,295
k206	114	24	3426.23	2	13	1	2	2	1,779,277,912
k207	84	42	4629.73	4	11	1	4	4	2,436,037,050
k207	114	22	3143.96	2	12	1	2	2	1,654,265,593
k208	84	44	4840.02	4	11	2	2	0	2,393,077,377
k209	114	56	7933.44	4	14	2	2	0	3,796,628,270
k401	84	28	3108.12	2	15	1	2	2	1,530,052,003
k401	84	28	3108.12	2	15	1	2	2	1,530,052,003
k402	49	86	6020.15	6	15	1	6	4	2,752,424,174
k403	84	50	5500.36	4	13	1	4	2	2,668,546,560
k404	49	52	3978.85	4	13	1	4	0	1,875,654,411
k404	49	76	4744.87	6	13	1	6	2	2,236,760,960
k405	49	63	4773.97	5	13	1	5	2	2,290,130,432
k405	84	52	5544.6	4	14	1	4	2	2,659,810,848
k406	84	56	6167.54	4	15	1	4	4	3,091,669,805
k407	84	22	2420.01	2	11	1	2	0	1,239,123,585
k407	84	50	5502.36	4	13	1	4	2	2,817,386,726

The table4 is the part of established database by using data analysis method.

4. DEVELOPMENT OF COST MODEL

4.1 Nearest-neighbor retrieval algorithm

As explained above, nearest-neighbor retrieval method is one of retrieval method in case-based reasoning which chooses the most similar case when the weighted sum of its features that match the current case is greater than other case in the case base. This research calls number of household, GFA, combination of household, floor, number of elevator, number of household per elevator and number of piloti household as attribute, so that there are 7 attributes in the database. To retrieve a similar case, it needs to calculate attribute similarity, attribute weight and case similarity. The figure5 represents nearest neighbor retrieval method algorithm.

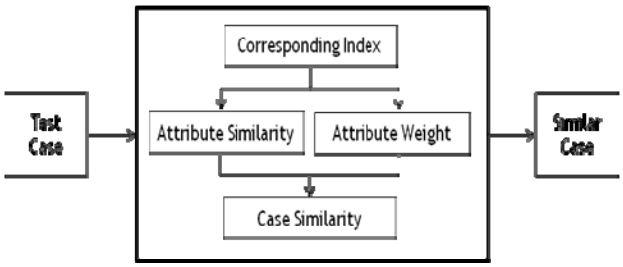

Fig. 5. Nearest-neighbor retrieval algorithm

4.2 Attribute similarity

A case is described by a set of attributes that characterize each element of the problem. The attributes of a case have their own data types. Data types can be categorized, according to their data characteristics, mainly into qualitative data and quantitative data. Similarity of qualitative data can be expressed by nominal scale. If text in test case appears to be exactly the same as text in database, then similarity S=1, or else similarity S=0. In case of quantitative data, its similarity is calculated by using ratio scale. Because this research has only quantitative data type, it needs tool to calculate similarity. The following formula (1) suggests calculating similarity.

$$f_{AS}(x) = \begin{cases} \dfrac{Min(AV_{Test\_case}, AV_{Retrieved\_case})}{Max(AV_{Test\_case}, AV_{Retrieved\_case})} & if, f_{AS}(x) \geq MCAS \\ 0 & if, f_{AS}(x) < MCAS \end{cases}$$

Where,
- $f_{AS}(x)$ = *Funtion of Attribute Similarity*
- AV_{Test_case} = *Attribute Value of Test_case*
- $AV_{retrieve\_case}$ = *Attribute Value of Retrieved_case*
- $MCAS$ = *Minimum Criteria for scoring Attribute Similarity*

4.3 Attribute weight

The major disadvantage of nearest neighbor retrieval method is to determine attribute weight. Attribute weights play an important role in CBR because the overall error obtained in CBR is a function of attribute weights (Dogan et al. 2008).

Some techniques are used to assess the weight. One of them is method using domain expert. The expert manually chooses the relevant attributes and indicates their relative importance for similarity assessment. However, this can be difficult and unstable because it is difficult to find the right expert who is knowledgeable about these issues and because expert opinion is subjective(Dogan et al. 2006). Second is technique by gradient descent method. Random cases are selected from the input case base, and the cases that are most similar to them are found, and then repeat this process many times. It uses gradient to decide the direction for search during the process. This process stops only when all gradients become zero. This research adopt genetic algorithm as a technique for determining attribute weight. Genetic algorithm provides optimum solution by using the "survival of the fittest" theory.

The total cost is derived from attribute multiplied by weight. Total cost can be expressed by following formula (2).

$$C_i = X_{i1} \times W_1 + X_{i2} \times W_2 + \cdots + X_{ij} \times W_j$$

where,

C_i = total cost of case i

X_{ij} = value of attribute j in case i

W_j = weight value of attribute j (applied to all cases)

The formula (2) can also be changed into matrix.

(3)

The W_j is optimum weight value for attribute j. However, because it is impossible to find the satisfied value for all cases in database, the general weight value should be calculated. The object to be optimized is difference between right clause and left clause.

For this research, GA software, Evolver was used to find the optimum attributes weight. The optimized object is sum of difference between right clause and left clause of formula (3). The table4 below shows the attribute weight which calculated by using Evolver.

Table 4. Result of Attribute Weight

Area type	W_1	W_2	W_3	W_4	W_5	W_6	W_7
49 ㎡	0.003	0.554	0.007	0.000	0.198	0.155	0.083
59 ㎡	0.002	0.715	0.096	0.080	0.055	0.000	0.052
84 ㎡	0.000	0.812	0.000	0.000	0.049	0.090	0.049

where,

W_1 = Weight value for # of household

W_2 = Weight value for Gross floor area(GFA)

W_3 = Weight value for Combination of household

W_4 = Weight value for # of household per a elevator

W_5 = Weight value for # of piloti household

W_6 = Weight value for Floor

W_7 = Weight value for # of elevator

4.4 Case Similarity

To calculate the case similarity is the last step for retrieving in nearest neighbor method. Through the 2 step of calculating attribute similarity and weight, it is relatively easy to measure case similarity. Case similarity can be calculated from the following formula (4).

$$f_{CS}(x) = \frac{\sum_{i=1}^{n} (f_{AS_i} \times f_{AW_i})}{\sum_{i=1}^{n} (f_{AW_i})}, \quad (\text{n} = \textit{the Number of Attribute})$$

Where, f_{CS} = Funtion of Case Similarity

f_{AS} = Funtion of Attribute Similarity

f_{AW} = Funtion of Attribute Weight

The table5 shows the result of calculating case similarity. The second column from left is the case similarity. Similar case can be retrieved based on the case similarity score. In table5, k711 data has the highest case similarity score.

4.5 Reuse cost information

In this research, because it is impossible to extract the same case with test case, the retrieved case can`t be used just as it is. It must be undergone the process of adaption. Once similar case is retrieved, the cost information becomes available. Estimator does not use total cost of retrieved case but use a unit cost per area(㎡). The new case s102 has 3091.49 ㎡ of gross floor area, and the retrieved case k711 has 3103.31 ㎡. The total cost of k711 is 1,669,328,023 won, so that unit cost per area(㎡) is 537,918 won. To validate the accuracy of cost model, 537,918 won of unit cost per area(㎡) is multiplied by 3091.49 ㎡. The estimated cost is 1,662,969,826 won. In comparison with the estimated cost, the actual cost for s102 is 1,638,576,786 won, which brings about 1.47% of error.

5. CONCLUSION

Because market change and severe competition become more serious in the construction industry, cost is considered to be critical success factor of construction projects. Particularly, cost at early stage influences more than that of other stages. Therefore, the purpose of this research is to suggest the cost model for early stage such as conceptual stage. The cost model is developed based on case-based reasoning(CBR). The different aspect of CBR compared with other Artificial Intelligence(AI) technique is that CBR utilizes the specific knowledge of previously experienced problem situations. Nearest neighbor retrieve method is used to choose the similar case. To retrieve it, attribute similarity, attribute weights and case similarity should to be calculated. Major weakness of nearest neighbor retrieval method is the difficulty to measure attribute weight. Genetic algorithm is search algorithms based on the mechanics of natural selection and natural genetics. In this research, attribute weight can be calculated by using genetic algorithm. If the similar case is selected based on nearest neighbor retrieval method, estimator can predict the total cost of new case. In order to validate the accuracy of cost model, this research uses data which are not included in database. As a result of validation, the error is 1.47%. However, this research limited only one data for validation, so that much more data is needed for validation.

Table 5. Result of case similarity

unit	area(㎡)	# of household	Attribute Similarity	GFA	Attribute Similarity	Com. Of househ	Attribute Similarity	# of household per an Elev.	Attribute Similarity	# of piloti household	Attribute Similarity	Floor	Attribute Similarity	# of Elev.	Attribute Similarity	Case Similarity	Total Cost
s102	84	28		3091.49		2		2		2		15		1			
k102	84	48	0.57143	5065.7	0.549	4	-	4	-	4	0.500	12	0.667	1	1.000	3.287	2,737,060,360
k103	84	42	0.66667	4629.73	0.649	4	-	4	-	4	0.500	11	0.556	1	1.000	3.371	2,616,869,987
k201	84	24	0.90476	2638.56	0.897	2	1.000	2	1.000	0	0.500	12	0.667	1	1.000	5.968	1,514,971,033
k202	84	48	0.52381	5277.12	0.501	4	-	2	1.000	0	0.500	12	0.667	2	-	3.191	2,714,112,388
k203	84	48	0.52381	5277.12	0.501	4	-	2	1.000	0	0.500	12	0.667	2	-	3.191	2,695,165,977
k205	84	40	0.71429	4448.13	0.690	4	-	4	-	4	0.500	11	0.556	1	1.000	3.460	2,458,138,750
k206	84	48	0.52381	5320.07	0.491	4	-	4	-	4	0.500	13	0.778	1	1.000	3.293	2,831,838,527
k207	84	42	0.66667	4629.73	0.649	4	-	4	-	4	0.500	11	0.556	1	1.000	3.371	2,496,937,976
k208	84	44	0.61905	4840.02	0.601	4	-	2	1.000	0	0.500	11	0.556	2	-	3.275	2,452,904,312
k401	84	28	1.00000	3108.12	0.996	2	1.000	2	1.000	2	1.000	15	1.000	1	1.000	6.996	1,568,303,303
k401	84	28	1.00000	3108.12	0.996	2	1.000	2	1.000	2	1.000	15	1.000	1	1.000	6.996	1,568,303,303
k403	84	50	0.47619	5500.36	0.450	4	-	4	-	2	1.000	13	0.778	1	1.000	3.704	2,735,260,224
k405	84	52	0.42857	5544.6	0.440	4	-	4	-	2	1.000	14	0.889	1	1.000	3.757	2,726,306,119
k406	84	56	0.33333	6167.54	0.298	4	-	4	-	4	0.500	15	1.000	1	1.000	3.131	3,168,961,550
k407	84	22	0.85714	2420.01	0.847	2	1.000	2	1.000	0	0.500	11	0.556	1	1.000	5.759	1,270,101,675
k407	84	50	0.47619	5502.36	0.450	4	-	4	-	2	1.000	13	0.778	1	1.000	3.704	2,887,821,394
k409	84	44	0.61905	4889.93	0.589	4	-	2	1.000	2	1.000	12	0.667	2	-	3.875	2,580,250,294
k506	84	56	0.33333	6146.43	0.303	4	-	4	-	4	0.500	15	1.000	1	1.000	3.136	3,296,181,716
k507	84	54	0.38095	5932.77	0.351	4	-	2	1.000	0	0.500	14	0.889	2	-	3.121	2,978,359,316
k508	84	54	0.38095	5932.77	0.351	4	-	2	1.000	0	0.500	14	0.889	2	-	3.121	2,976,210,097
k509	84	44	0.61905	4907.92	0.585	4	-	2	1.000	4	0.500	12	0.667	2	-	3.371	2,641,327,716
k601	84	54	0.38095	5950.35	0.347	4	-	4	-	2	1.000	15	1.000	1	1.000	3.728	3,140,993,837
k602	84	52	0.42857	5750.06	0.393	4	-	4	-	4	0.500	15	1.000	1	1.000	3.322	3,135,477,065
k603	84	52	0.42857	5750.06	0.393	4	-	4	-	4	0.500	15	1.000	1	1.000	3.322	3,145,639,403
k604	84	52	0.42857	5750.06	0.393	4	-	4	-	4	0.500	15	1.000	1	1.000	3.322	3,137,213,687
k608	84	20	0.80952	2200.62	0.797	2	1.000	2	1.000	0	0.500	10	0.444	1	1.000	5.551	1,217,032,857
k610	84	28	1.00000	3108.11	0.996	2	1.000	2	1.000	2	1.000	15	1.000	1	1.000	6.996	1,693,729,494
k708	84	56	0.33333	6189.09	0.293	4	-	4	-	4	0.500	15	1.000	1	1.000	3.126	3,103,753,684
k709	84	56	0.33333	6189.09	0.293	4	-	4	-	4	0.500	15	1.000	1	1.000	3.126	3,127,108,177
k710	84	52	0.42857	5753.22	0.392	4	-	4	-	4	0.500	14	0.889	1	1.000	3.210	3,080,226,903
k711	84	28	1.00000	3103.31	0.997	2	1.000	2	1.000	2	1.000	15	1.000	1	1.000	6.997	1,669,328,023
k801	84	50	0.47619	5551.9	0.438	4	-	2	1.000	4	0.500	14	0.889	2	-	3.303	2,903,449,379
k803	84	44	0.61905	4863.04	0.596	4	-	4	-	4	0.500	12	0.667	1	1.000	3.381	2,633,151,864
k804	84	44	0.61905	4860.72	0.596	4	-	4	-	4	0.500	12	0.667	1	1.000	3.382	2,751,251,273
k805	84	44	0.61905	4897.56	0.588	4	-	2	1.000	4	0.500	12	0.667	2	-	3.373	2,685,002,495
k808	84	22	0.85714	2101.45	0.774	4	-	2	1.000	0	0.500	6	-	2	-	3.131	1,431,223,118
k809	84	22	0.85714	2496.54	0.864	4	-	2	1.000	4	0.500	8	0.222	2	-	3.444	1,608,584,306
k901	84	22	0.85714	2453.89	0.854	2	1.000	2	1.000	4	0.500	12	0.667	1	1.000	5.878	1,242,159,574
k903	84	22	0.85714	2453.89	0.854	2	1.000	2	1.000	4	0.500	12	0.667	1	1.000	5.878	1,184,512,827
k910	84	56	0.33333	6163.58	0.299	4	-	4	-	4	0.500	15	1.000	1	1.000	3.132	3,370,266,054

ACKNOWLEDGEMENT

This research was supported by a grant (06-A03) from Construction Technology Innovation Program funded by Ministry of Land, Transport and Maritime Affairs of Korean government.

REFERENCES

[1] Nie-Jia Yau & Jyh-Bin Yang, Case-Based Reasoning in Construction Management, Computer-Aided Civil and Infrastructure Engineering, 13(1998),143-150

[2] Ian Watson & Farhi Marir, Case-Based Reasoning: A Review, The Knowledge Engineering Review, Vol.9 No.4, 1994

[3] Agnar Aamodt & Enric Plaza, Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches, AI Communications. IOS Press, Vol.7:1(1994), pp.39-59

[4] Sevgi Zeynep Dogan, David Arditi & H.Murat Gunaydin, Determining Attribute Weights in a CBR Model for Early Cost Prediction of Structural Systems, Journal of Construction Engineering and Management, Vol.132 No.10, 2006, pp.1092-1098

[5] Sevgi Zeynep Dogan, David Arditi & H.Murat Gunaydin, Using Decision Trees for Determing Attribute Weights in a Case-Based Model of Early Cost Prediction,

Journal of Construction Engineering and Management, Vol.134 No.2, 2008, pp.146-152
[6] SANHAR K. PAL & SIMON C. K. SHIU, Foundation of Soft Case-Based Reasoning, Wiley-Interscience, 15-17
[7] Ian Watson, Applying Case-Based Reasoning, Morgan Kaufmann Publishers, 23-33
[8] David E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, 10-14.