Military Facility Cost Estimation System Using Case-Based Reasoning in Korea

Sae-Hyun Ji[1]; Moonseo Park, M.ASCE[2]; Hyun-Soo Lee, M.ASCE[3]; Joseph Ahn[4]; Namho Kim[5]; and Bosik Son[6]

Abstract: Numerous cost estimations are made repetitively in the initial stages of construction projects in response to ongoing scope changes and often need to be recalculated frequently. In practice, the square foot method, considered an effective method for time-saving, is widely used. However, this method requires a great amount of effort to calculate a unit price and does not consider the uniqueness of each case. Thus, the use of the square foot method could bring about unwanted consequences. For example, in the case of military projects in Korea, significant differences have been reported between estimations made using this method and the actual costs. In an effort to deal with this challenging issue, this research develops a military facility cost estimation (MilFaCE) system, based on case-based reasoning (CBR), using case data from 422 construction projects at 16 military facilities. Based on system validation experiments involving 10 military officers (engineers), the effectiveness of the system in terms of estimation accuracy and user-friendliness is confirmed. Consequently, this research can be a CBR application example of construction cost estimation and a basis for further research into the development of cost estimate systems.
DOI: 10.1061/(ASCE)CP.1943-5487.0000082.

CE Database subject headings: Construction costs; Databases; Korea, South; Military engineering; Estimation; Computation.

Author keywords: Cost; Estimate; Case-based reasoning; Database.

Introduction

Efficient expenditures for military construction projects is critical because these projects use a significant portion of a total military budget (over 6.5% in Korea in 2009). Currently in Korea, budgeting for military construction is based on the Standard and Guide of the National Defense Budget Organization of Korea, hereafter called "the standard" (Korea Ministry of National Defense 2008), which recommends the square foot method by facility type as the standard cost-estimating method. This traditional parametric method has been considered an effective one to reduce the time spent on an estimate. However, in practice, it requires promulgation of new unit prices, which are generally based on the mean of data, periodic analysis of new projects (e.g., quarterly or annually), and historical adjustment of analyzed data, leading to a time-consuming process for computing unit prices. Furthermore, because there are significant differences between estimations based on this standard and actual expenditures, the aim of cost-planning in the preconstruction phase is oriented toward "designing to a cost" in the case of insufficient budgets. As a result, owners cannot help making many changes resulting from cost overruns. This situation can become worse when unit prices are not updated in a timely manner. Actually, neither has the standard been updated, nor were new analyses provided; thus the unit prices in the standard from 1999–2006 remained the same. Good accuracy is hardly expected in this situation. Consequently, the Korea Ministry of National Defense has not been able to implement all of their intended projects, and the Korean Army could not accomplish 15% of their plans for the construction of military quarters in 2008 because of a budget shortage.

Based on interviews, it was determined that most military officers (engineers) engaging in budgeting for construction projects would like an advanced cost-estimation system to support their decision-making. In an effort to deal with this challenging issue, this research develops a cost-estimation system which utilizes a case-based reasoning (CBR) method with genetic algorithms [military facility cost estimation (MilFaCE) system]. To achieve the research goal, first the research scope was defined as the initial-stage cost estimation. In order to determine the trends of CBR-applied approaches, previous research was reviewed. Then, data were collected, and a database was constructed with the assistance of the Korea Ministry of National Defense. Having constructed a cost database, we developed a cost model with cost-estimation methodologies which were mapped into the estimation process. Then, a data model was developed for MilFaCE. Finally, the effectiveness of the system was validated in terms of both user convenience and estimation accuracy. The research outcome presents potential for commercialization of a CBR-based cost-estimation system in the construction industry by providing an advanced CBR estimation tool to industry practitioners. Furthermore, this research could be a basic model of process- or methodology-building for future research about system development in the construction industry and CBR system development for other industries.

[1]Ph.D. Candidate, Dept. of Architecture, Seoul National Univ., San 56-1 Shinrim-dong, Seoul, Korea. E-mail: oldclock@snu.ac.kr

[2]Associate Professor, Dept. of Architecture, Seoul National Univ., San 56-1 Shinrim-dong, Seoul, Korea (corresponding author). E-mail: mspark@snu.ac.kr

[3]Professor, Dept. of Architecture, Seoul National Univ., San 56-1 Shinrim-dong, Seoul, Korea. E-mail: hyunslee@snu.ac.kr

[4]Research Student, Dept. of Architecture, Seoul National Univ., San 56-1 Shinrim-dong, Seoul, Korea. E-mail: josephahn@snu.ac.kr

[5]Programmer, °M.S.E., Doalltech, 1580 Sangam-dong, Mapo-gu, Seoul, Korea. E-mail: namo@doalltech.com

[6]Assistant Professor, Dept. of Architectural Eng., South Seoul Univ., Seobuk, Chonan, Korea. E-mail: bsson@nsu.ac.kr

Note. This manuscript was submitted on January 22, 2010; approved on July 30, 2010; published online on August 3, 2010. Discussion period open until October 1, 2011; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Computing in Civil Engineering***, Vol. 25, No. 3, May 1, 2011. ©ASCE, ISSN 0887-3801/2011/3-218–231/$25.00.

Preliminary Research

Traditionally, cost estimates are based on the estimator's experience, imaginative abilities, and a wide range of assumptions including appraisals of previously conducted projects that are similar in scope (Jarde and Alkass 2007). In practice, parametric cost estimation (e.g., cost per square foot) is one of the most common methods in the initial project phase and uses either historical cost data or cost books to obtain an estimate of a building's cost per square foot (Karshenas 2005; Kirkham 2007). This method does not require detailed information of a project and is thus relatively less time-consuming for estimating approximate costs (AACE 1999). Recently, Hegazy and Ayed (1998), Lowe et al. (2006), and Soutos and Lowe (2005) developed a cost model that utilizes a multiple regression equation based on building cost data. However, this method cannot overcome the problem related to the complex interactions among factors, which have a negative impact on estimate accuracy and employment of regression analysis (Graza and Rouhana 1995; Ji et al. 2009). Also, limitations of regression analysis exist in defining mathematical formulas for the cost function from selected past projects (Creese and Li 1995) and in the dependence of its robustness on historical data of similar projects (Sonmez 2008). A neural network approach to estimating cost has also been triedt which can be beneficial when intuitive judgment is involved, or when data patterns become irregular (Hegazy and Ayed 1998). However, it takes time to determine the network factors that best fit the application, and the process of this method is regarded as a "black box" (Bode 1998; Smith and Mason 1996). In this context, the case-based reasoning method is being highlighted as a decision-making tool for the construction industry. A similarity measurement in CBR is particularly important during the retrieval process and is an attempt at handling the relationships between the relevant objects associated with the cases (Pal and Shiu 2004). Generally, the measurement is based on the sum of similarities of the case representative attributes. For these computations, an optimized assignment of attribute weight values is required. This research adopts genetic algorithms as an optimization method. The following subsections will discuss CBR and genetic algorithms.

Case-Based Reasoning

Case-based reasoning originates from the 1977 work of cognitive scientists Schank and Abelson. They proposed that our general knowledge about a situation is recorded in the brain as a script that allows us to set up expectations and perform inferences (Watson 1997). CBR is regarded as a plausible high-level method for cognitive processing. It focuses on problems such as how people learn a new skill and how humans generate hypotheses about a new situation based on their past experiences (Pal and Shiu 2004). The processes of CBR can be seen as a reflection of a particular type of human reasoning in which people generally solve problems encountered with an equivalent of CBR (Kolodner 1993). Recently, many studies have been conducted in the construction industry related to CBR for cost estimation purposes (Yau and Yang 1998; Karshenas and Tse 2002; An et al. 2007; Yi et al. 2004; Chou 2009), international market selection (Ozorhon et al. 2006), decision-making support (Chua et al. 2001; Morcous et al. 2002; Chua and Loh 2006), planning and management (Yau and Yang 1998; Tah et al. 1998), scheduling (Ryu et al. 2007), and predicting the outcome of litigation (Arditi and Tokdemir 1999). Although these show the potential of application of the CBR system to the construction domain, the system only considers the direct cost of buildings (Yau and Yang 1998) and does not provide project-related information for the selected cases (Chou 2009). The CBR method uses conceptually straightforward approaches to approximate real-valued or discrete-valued target functions. Its learning algorithms consist of simple storing processes. When a new query case is encountered, a set of similar, related cases is retrieved from memory and is used to classify a new query case (Mitchell 1997). In this context, establishing the computation of similarity can be a key issue for the whole CBR process. In the literature, there are several measurement methods (Burkhard 2001; Ozorhon et al. 2006; An et al. 2007; Ryu et al. 2007; Qian et al.2009; Chou 2009) that often lack an explanation and are incomputable when the target case exists outside the case-base range. Thus, this research employs the similarity-measuring formula using distance of the location between objects in Euclidean space as below (Ji et al. 2009). SIM (x_i, x_j) = degree of similarity between x_i and x_j, and DIS (x_i, x_j) = weighted distance between the two cases x_i and x_j, where $a_r(x)$ = value of the rth attributes of case x, and w_r = weight of the attributes of the case. All the attributes' values are converted to new scores of 0 to 1, applied by the probability density function; when the square root of the sum of squares of the weight values is assigned as 1 ($\sum w_r^2 = 1$), then the range of the weighted distance of the two cases can be standardized by [0, 1]

$$\mathrm{SIM}(x_i, x_j) = 1 - \mathrm{DIS}(x_i, x_j) = 1 - \sqrt{\sum_{r=1}^{n} w_r^2 [a_r(x_i) - a_r(x_j)]^2} \tag{1}$$

Genetic Algorithms

Genetic algorithms (GAs) are search algorithms based on the mechanics of natural selection and genetics. GAs are iterative procedures that maintain a population of candidate solutions to optimize a fitness function. Having been established as a valid strategy for problems requiring efficient and effective searching, GAs have been widely used for many applications in business, scientific, and engineering circles, as they provide simplicity in computation and are powerful in their search for improvement (Goldberg 2006). GAs are used to search a space of candidate hypotheses to identify the best hypothesis. The best hypothesis is defined as the one that is the optimized value for the predefined numerical measure at hand, which is called hypothesis fitness (Mitchell 1997). This research uses the hypothesis function fitness suggested by Ji et al. (2009) which seeks the optimal value of ω_i to minimize the sum of the square root of the distance (i.e., Euclidian distance) between each side of the equation shown as Eq. (2). Let C_j, ω_i, and X_i = cost of the jth case project, the weight value of the ith attribute, and the ith attribute value of the jth case, respectively

$$\min \sum_{n=1}^{j} \sqrt{D_n^2},$$
$$\text{s.t.} \begin{pmatrix} C_1 \\ \vdots \\ C_j \end{pmatrix} - \begin{pmatrix} X_{11} & \dots & X_{1i} \\ \vdots & \ddots & \vdots \\ X_{j1} & \dots & X_{ji} \end{pmatrix} \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_i \end{pmatrix} = \begin{pmatrix} D_1 \\ \vdots \\ D_j \end{pmatrix} \tag{2}$$

Cost Model Development

Cost data in Korea generally consist of four components: the cost of preliminary work, site work, buildings, and indirect costs. Based on data from 422 cases (Table 1), the cost ratios of each component are analyzed to quantify their percentages. As a result, it is discovered that the average cost ratio of preliminary work is 0.42%, the average cost ratio of site work is 7.45%, the ratio of buildings is 68.41%, and the ratio of indirect costs is 23.72%. Apparently, the sum of building costs and indirect costs on average comprises

J. Comput. Civ. Eng. 2011.25:218-231.

Table 1. Data Profile and Database Configuration

Information	Number of cases	Components/attributes
Cost information	—	ID, cost of preliminary works, building cost, site work cost, indirect cost
Project general information	—	ID, facility category, facility name, Army/Air Force/Navy, year, region, name of building, data input
Site work information	—	ID, site area, building-to-land ratio, shape of site, retaining wall, capacity of septic tank, rocks at foundation, reinforcement of an incline, area of drill ground
Building information	422	
Barracks	205	
Quarters	151	ID, magnitude, number of beds (NB), shape of roof (SR), structure type (ST), number of floors (NF), envelope materials (EM), gross floor area (GFA), unit floor area (UFA), quarter area ratio (QR), number of underground floors (NUF), pit, office area ratio (OR), dining area ratio (DR), bathhouse area ratio (BR), pile foundation (PF), type of heating (TH), air conditioning (AC)
Mess halls	35	ID, magnitude, GFA, UFA, EM, DR, SR, NF, NUF, pit, ST, PF, seating capacity (SC)
Bathhouses	19	ID, magnitude, GFA, UFA, EM, SR, NF, NUF, ST, PF, pit, accommodation capacity (ACC)
Maintenance and arsenals	105	
Warehouses	53	ID, magnitude, GFA, UFA, EM, SR, NF, NUF, ST, PF, pit, AC, racks (RA), purpose (PP), building height (BH)
Magazines and arsenals	26	ID, GFA, UFA, EM, SR, NUF, ST, PF, pit, RA, BH
Maintenance shops	15	ID, magnitude, GFA, UFA, EM, SR, NF, NUF, ST, PF, pit, work area ratio (WR), height of work area (HW)
Car shades	11	ID, magnitude, GFA, UFA, EM, SR, ST, PF, BH
Supporting facilities	49	
Office buildings	24	ID, magnitude, GFA, UFA, EM, SR, NF, NUF, ST, PF, pit, PP, other purpose area ratio (OPR)
Guard houses	4	ID, magnitude, GFA, UFA, EM, SR, NF, NUF, ST, PF, pit
Gyms	9	ID, GFA, UFA, EM, SR, NF, NUF, ST, PF, pit, AC, TH, BH
Welfare facilities	6	ID, magnitude, GFA, UFA, EM, SR, NF, NUF, ST, PF, pit, AC, BR, TH, swimming area ratio (SR), lodging area ratio (LR)
Interview houses	6	ID, magnitude, GFA, UFA, EM, SR, NF, NUF, ST, PF, pit, PP, TH, AC, interview area ratio (IR)
Living quarters	63	
Apartments	3	ID, GFA, UFA, SR, NF, NUF, ST, PF, pit, area type (AT), number of households (NH), number of unit floor households (NUH), hallway type (HT), unit plan type (UT)
Bachelor officers' quarters	50	ID, GFA, UFA, EM, SR, NF, NUF, ST, PF, pit, NH
Single-family houses	8	ID, GFA, UFA, EM, SR, NF, NUF, ST, PF, pit, NH
Row houses	2	ID, GFA, UFA, EM, SR, NF, NUF, ST, PF, pit, NH

over 90% of the total cost. One interesting fact is that indirect costs of public projects in Korea (e.g., military construction projects) should be determined by legislation according to building types and scale (i.e., cost). Therefore, it is expected to have a certain ratio of indirect cost over the total cost. With this in mind, this research adopts CBR as the estimation method for building cost, and the average ratios over all facility types as the cost estimation method of preliminaries, site work, and indirect cost. In addition, a regression analysis-based linear equation is used to adjust historical escalation rates which are utilized for the Korean construction cost index.

Data Acquisition and Analysis

According to the standard (Korea Ministry of National Defense 20082008), military facilities are classified into four building types based on their main functions: barracks, maintenance and arsenal facilities, supporting facilities, and quarters. These are again broken down into unit facilities as follows:

- Barracks: quarters, mess halls, and bathhouses;
- Maintenance and arsenal facilities: warehouses, magazines and arsenals, maintenance shops, and car shades;
- Supporting facilities: office buildings, guard houses, gyms, welfare facilities, and interview houses; and
- Living quarters: apartments, bachelor officers' quarters, single-family houses, and row houses.

Data for 422 implemented and planned military facility construction projects from 2004—2009 were collected (Table 1). For precise budget estimations, public owners in Korea make priced bills of quantities that contain the total expenditure for all inputs (e.g., labor and materials) paid to contractors. The government strictly regulates the pricing methods for these inputs. Therefore, public owners regard the total cost based on priced bills of quantities as the only standard for precise budgeting. This is what we collected and analyzed for development of this research. Based on this cost, construction firms determine a total fixed price for bidding which is lower than the owners' price and contract with this price (contractor suggested price). Thus, issues related to performance of a project, such as a schedule delay or cost-overrun caused by the owner's change-order, belong to the contractor's perspective. From the public owner's budgeting perspective, they did not consider the contractor's suggested price or cost of project completion. Because the standards of the suggested price differ according to contractors, there is a differential between the owner-expected cost and the contractor-suggested price. The owners believe that there is a government-regulated method on which outcomes are based (i.e., priced bills of quantities are the basis for the detailed project budget). Therefore, this research was developed using the detailed project budget, so that all of the data can be regarded by owners as actual data.

J. Comput. Civ. Eng. 2011.25:218-231.

A database was developed to have four information categories: project general information (GI), cost information (CI), site work information (SI), and building information (BI). The building information (BI) tables are formatted differently according to facility types. Attributes in the database tables are organized according to the purpose or function of a building. However, use of all attributes for BI is inefficient (BI has 35 kinds of attribute). Because other information tables (CI, PI, and SI) are each organized by the same formats, they can be represented by a relatively small number of attributes in common (e.g., CI has five, PI has seven, and SI has nine attributes). Each of them can be searched and extracted using its own identification name. Furthermore, these are considered to be developed in a relational data model which can provide what the users want to know after transformation. With this in mind, we conducted a comprehensive analysis of drawings and subsequently extracted information. Specifically, project general information has 13 attributes, cost information has 17, site work information has nine, and building information has 9–18 attributes, according to facility type (Table 1). The appropriateness of these attributes has been confirmed by military officers (Army, Navy, Air Force, and Marine Corps) who have been working for the Department of Civil Engineering as cost estimators for over ten years.

Cost Model Scheme

In the context of the aforesaid discussions, the cost model and its process are organized by each component of cost data. As diagrammed in Fig. 1, the model consists of three components: the CBR module, the supplement module, and the adjustment module. The CBR module utilizes the case-based reasoning method for building cost estimation. The other modules work with the results of the CBR module. On closer examination, the process initiates with the selection of facility type and the inputting of information about the facility's attributes. The system then begins to compute similarity scores for each case in the database. When these cases are arrayed in ascending order, they have cost values that are normalized to the standard year (2008) with respect to escalation, using the Korean construction cost index. This index classifies 16 types of facilities that are officially announced every month. From these cases, plural similar cases [K-nearest neighbors (K-NN)] would be extracted, and the average unit cost calculated based on these K-NN cases. Thereafter, building cost would be estimated. The cost ratios of site work, preliminary work, and indirect costs of the selected facility type would be simultaneously loaded from the database to the supplement module. These cost ratios for a pertinent facility type are used for estimating preliminary work cost, site work cost, and indirect cost by being multiplied by the estimated building cost. Finally, the estimated cost would be reported after being adjusted for regional differences and escalation in the adjustment module (Fig. 1).Because of Korea's relatively small territory, the method of regional adjustment is defined, related to working conditions, by legislation and suggested marginal guidelines regarding labor cost increments.

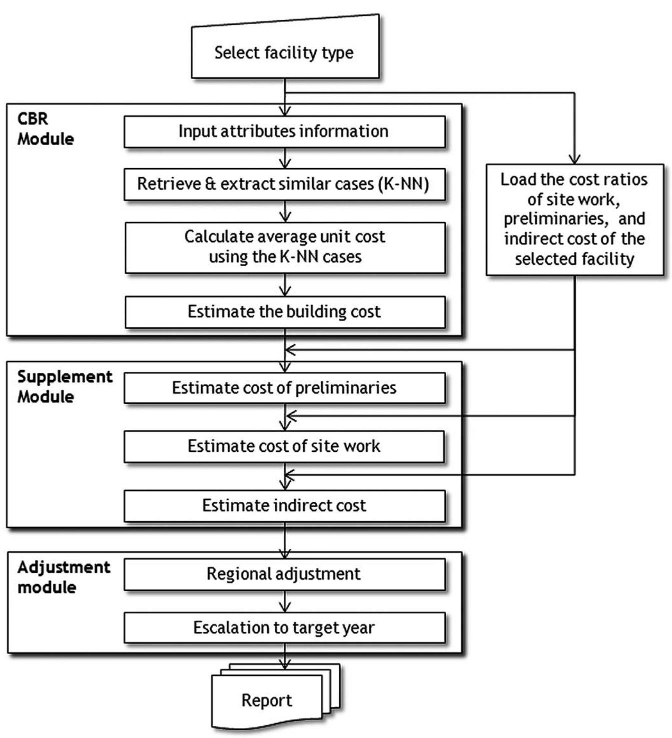


Fig. 1. Cost model scheme

However, when midterm planning for military construction is updated every year, budgeting for new projects that can be executed in five years takes precedence. Consequently, it is necessary that budgeting take into account the future value.

System Development

Military facility cost estimation is developed to be a stand-alone tool which is interconnected via the relational database system. Basically, the system is developed using C# language and Microsoft *Access* program. The MilFaCE system starts an operation when a user chooses a facility type and completes inputting building information. The system then begins iteratively computing the similarity scores at the pertinent sector of case library. Basic statistical conversion functions are included in the system which are necessary for similarity computation. One of the technical challenges we encountered for system development was statistical conversion. All values should be converted to [0, 1], based on statistical functions, to compute case similarity [Eq. (1)] and to assign attribute weight values using genetic algorithms [Eq. (2)]. To solve this issue, the data are converted twice using statistical functions. One is the standardization that can change a raw value to a value in the standard normal distribution. The other is the probability density function that can calculate the cumulative probability at a given point of the standard normal distribution. The adoption of these functions might slow working speed of this system slightly because the conversions should have iterative precedence over similarity computations for every single case. If these repetitive calculations are done, the system will require the database, using database queries, to show all information for each case. This means that the MilFaCE system is configured to be connected by database queries between the operation unit and each independent case library in the database. For that reason, MilFaCE has many embodied database queries that can refer to various pertinent case libraries to create outputs in accordance with the inputted information. The system can get updated easily by adding new raw data without processing. This is the greatest advantage of the system. Thus, effort and time for data analysis and the possibility of error occurrence can be reduced.

Data Modeling

A data model is a diagram to represent groupings of data so that the reader can better understand the actual data (Hoberman 2005). The main objective of a data model is to provide flexible and timely access to the correct data source to enhance accessibility. Data modeling can be defined as the process of formalizing the data requirement using a conceptual modeling technique (Sanders 1995).

Estimators generally refer to project information and cost data when searching for similar cases. The system can provide the required correct information from a database in a timely manner. However, the data have many attributes. Cost information, in particular, must be processed (an escalation or arithmetic operation) before being presented in the system. Thus, defining database tables and clarifying their relationships should take precedence over system development. As shown in Fig. 2, the data model is developed using an entity relationship diagram which has 19 tables of information: cost information, project general information, site work information, and 16 tables of building information. All building and site work information tables have respective relationships with the project general information table; and the cost information and project general information tables are cross-referenced. The database system, which is under the data-loading capacity, is mainly aimed to retrieve information. Consequently, performance of the system might not be affected to a great degree by designation of primary a key (PK) and foreign key (FK). Because it is too complicated to visualize all of the existing relationships, only the PK is shown in Fig. 2. Therefore, the schema of the system is built in such a way that cost information, building information, and site work information tables are linked around the project information table. The identification (ID) of each table is PK, and its relationship to the PI table is linked as a one-to-one relationship that results in the creation of FK. Based on this data model, the CBR module of the system is developed to respond in real time when a case-retrieving query is required.

CBR Module

Both quantitatively and qualitatively, case storage is crucial for developing the CBR system. Because each time a new query case is encountered in the CBR system, its relationship to the previously stored examples is examined to assign a target function value for the new case (Mitchell 1997). In the qualitative respect, if the case were stored based on an analysis using a multifacet approach, this potentially useful data could support users' decision-making more accurately. A certain numbers of cases are necessary in the CBR method to determine attributes and assign attribute weight values. With this in mind, in applying the suggested CBR method, this research used six types of facilities that have over 15 cases and meaningful attributes in addition to gross floor area: quarters, mess halls, warehouses, magazines and arsenals, maintenance shops, office buildings, and bachelor officers' quarters. These buildings have their own attributes whose types can be classified into numerical or true/false or nominal attributes. Among these, numerical- and true/false-type attributes' information is used when applying the similarity function and weight assignment method. As summarized in Table 2, the attribute weight-values of each facility are assigned using these attributes and the GA-based weight assignment method. On closer examination, gross floor area has the highest score of weight-value in most instances, and the maintenance shops ranked second. In this respect, we employed the traditional square-foot method for the other kinds of buildings which are unsuitable for the CBR method: bathhouses, car shades, guard houses, gyms, welfare facilities, interview houses, apartments, single-family houses, and row houses. If users input the gross floor area, then cases in the database will array with their own information, and users can choose similar cases heuristically and use them to estimate the project cost.

Assuming that similar problems have similar solutions, retrieved cases can be used for solving a new problem. Generally, a past solution needs adjustment to fit the new situation and is called case adaptation. Case adaptation is the process of transforming a retrieved solution into a solution appropriate for the current problem. Thus, case adaptation (adjustment) is a crucial step to the CBR method because it adds intelligence to what would otherwise be simple pattern-matching (Pal and Shiu 2004). There are two points of view related to the adjustment methods of retrieved cases: (1) What kinds of methods would be used? (2) How many cases would be selected for solving the problem?

Regarding the first issue, this research applies an algorithm that calculates the mean value of the k-nearest neighbors rather than calculating the most similar value. Because the k-nearest neighbor algorithm is the most basic instance-based method for approximating a real-valued or discrete-valued target function. Also, the target function value for a new problem is estimated from the known value of the k-nearest neighbor cases (Mitchell 1997). This involves searching for the k nearest cases to the current input case, based on similarity measures, then selecting the class of the majority of these k cases as the retrieval case. Precisely, we propose a normalizing method of getting an independent variable (i.e., cost) divided by the most heavily weighted attribute, gross floor area (GFA). The second issue concerns how many cases would be selected to solve the problem. Particularly in an experience-oriented industry such as construction, knowledge and assessments of previous projects are essential in resolving reoccurring problems. Much research related to this issue has been conducted. Chua et al. (2001) used two sets of similar cases: one set focused on the subgoal of competition, and the other set focused on supporting bid decision-making. Ozorhon et al. (2006) assumed that a 70% similarity score is sufficient for the final prediction. An et al. (2007) applied the most similar case to estimate construction cost. Arditi and Tokdemir (1999) chose a threshold-cutoff similarity score of 75% to predict the outcome of construction litigation. Ahn et al. (2006) proposed an approach for CBR which applies a genetic algorithm as a simultaneous optimization algorithm for determining the number of cases in k-nearest neighbor. Their experiment was conducted with a k range of 1–10. Despite these attempts, it is rarely certain which method is the best. Hence, MilFaCE was developed to present all of its own special information with similarity scores in ascending order. In other words, the system gives users the option to select similar cases, according to their own experience-based decision.

Supplement Module

As discussed in preceding sections, the cost estimation method of the supplement module is based on cost ratios of corresponding facility types over building cost. The average percentages compared to the building cost (or the sum of building and site work cost) are analyzed and summarized in Table 3 according to facility types. Some data which have abnormal (extremely high or low) values have been excluded by heuristic decisions (mistakes are suspected because the cost ratios are expected to be similar owing to legislative control), and certain quantitatively insufficient types of buildings use the average value of the pertinent building group, i.e., guard houses, interview houses, single-family houses, and row houses. The MilFaCE system provides these cost ratios as default values; however, users can modify these values if necessary. The supplement module should be renewed at regular intervals by updating the cost ratios based on additional data analysis.

Adjustment Module

As described in the preceding section, the system produces an estimation outcome using present values. However, budgeting must carried out for a future project in which the actual implementation in the field will commence in five years. Hence, the estimation outcome must be escalated to adequately allow for increased costs at the planned date of commencement of construction. To overcome

J. Comput. Civ. Eng. 2011.25:218-231.

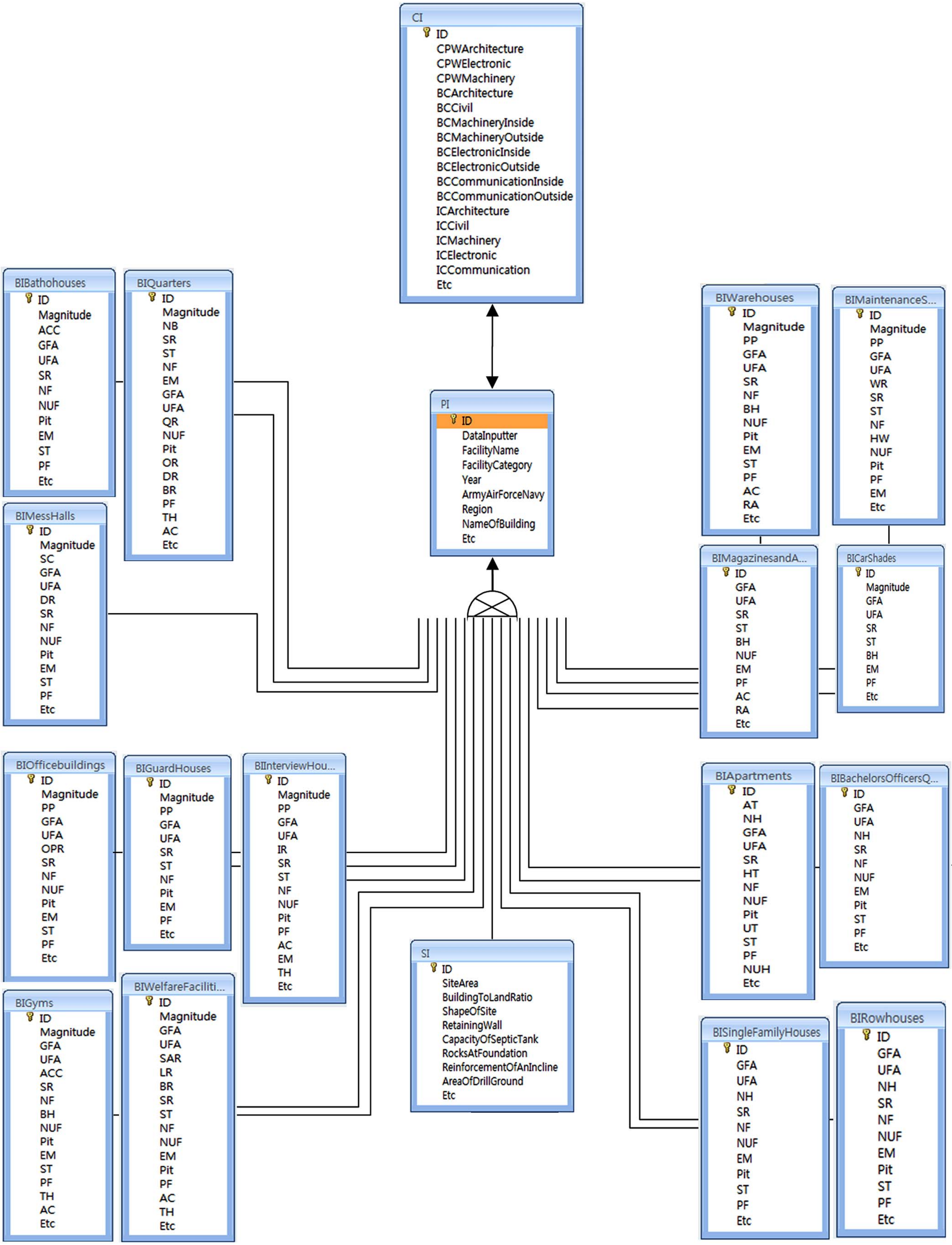


Fig. 2. Data model of MilFaCE system

Table 2. Weight Values Assigned by GA Method

Facilities	Weight values of attributes used																	
	GFA	UFA	NF	NB	NH	AC	NUF	PF	Pit	QR	OR	DR	SC	RA	WR	BH	HW	OPR
Barracks																		
Quarters	0.99871	0.00842	0.0306	0.02888	—	0	0	0	0.02718	3.3E–6	0	—	—	—	—	—	—	—
Mess halls	0.99757	0.00082	0.06952	—	—	—	—	0	0.00508	—	—	0	0.00061	—	—	—	—	—
Maintenance and arsenals																		
Warehouses	0.99985	0.01372	5.8E–05	—	—	0.00017		0.00041	0.00998	—	—	—	—	0		0		—
Maintenance shops	0.44638	0.86784	0.20200	—	—	—	0.08126	0.00071	0.00560	—	—	—	—		0.01256		0.00027	—
Supporting facilities																		
Office Buildings	0.98961	0.14107	0.02479	—	—	—	0.01202	2.1E–05	0.00016	—	—	—	—	—	—	—	—	0
Living quarters																		
Bachelor officers' quarters	0.62144	0.26327	0.05458	—	0.00088	—		0.05501	0.00481	—	—	—	—	—	—	—	—	—

Note: The sum of squares of all weight values for each facility is 1.0.

Table 3. Cost Ratios of Supplement Module

Facility	Average cost ratios		
	Preliminaries[a]	Site work[a]	Indirect cost[b]
Barracks			
Quarters	0.8%	14.9%	33.9%
Mess halls	0.6%	9.3%	33.2%
Bathhouses	0.5%	9.6%	36.1%
Maintenance and arsenals			
Warehouses	2.1%	15.8%	29.9%
Magazines and arsenals	0.0%	39.7%	33.4%
Maintenance shops	0.8%	13.7%	28.2%
Car shades	3.9%	48.9%	28.4%
Supporting facilities			
Office buildings	0.7%	7.1%	32.4%
Guard houses	1.5%	5.7%	29.3%
Gyms	0.9%	12.4%	33.3%
Welfare facilities	0.4%	12.7%	26.9%
Interview houses	1.5%	14.4%	41.0%
Living quarters			
Apartment	0.4%	8.0%	19.3%
Bachelor officers' houses	0.6%	8.0%	32.4%
Single-family houses	0.6%	8.%	32.5%
Row houses	0.6%	8.1%	32.5%

[a]Percentage compared to building cost.
[b]Percentage compared to sum of building and site work cost.

this transformation problem, this research derives an equation to convert this outcome to the future value using the Korean construction cost indices data. The indices, officially announced every month, are classified into 16 facility types and are designed to convert a past value to a present value. Basically, the indices are developed using the modified Laspeyres formula which is one of the general index-deriving formulas for calculating a normalized average of prices for a given class of goods or services in a given region during a given time-interval. Specifically, the indices are derived based on 105 items that represent the construction industry from January 2000 and are used for escalating cost estimates to the present value in the Korea construction domain. Based on these Korean construction cost indices data, the time-adjustment equation for producing future values is derived using simple regression analysis [Eq. (3)]. For increasing the robustness of these kinds of time-series analysis, this research utilized 118 indices of building construction from January 2000 to September 2009

$$\text{Cost Index}\,(\%) = 0.410 + 73.498 * [12(\text{year}_t - 2000) + \text{month}_{\text{ordinal}}] \tag{3}$$

Let year_t and $\text{month}_{\text{ordinal}}$ = target year and ordinal number of month of a prearranged project execution. This equation has 0.911 of adjusted R-square value which means the equation could be represented over 91% of the data.

Because military construction budgeting in Korea is updated on an annual basis, the system uses the annual average cost indices for time adjustment (Table 4).

Consequently, MilFaCE has a main window which has 17 respective facility windows, categorized into four groups. Each facility window is made up of two tabs: an input information and case retrieval tab and a computation and estimation report tab (Fig. 3 and 4). The former is used to search similar cases for estimation of a building cost, and the latter is used to report

J. Comput. Civ. Eng. 2011.25:218-231.

Table 4. Cost Indices for Time Adjustment

Year	Cost index (%; base year 2008)
2010	102.62
2011	106.64
2012	110.67
2013	114.70
2014	118.73
2015	149.75
2016	126.78

the cost estimation of a project (building cost, preliminary cost, site work cost, indirect cost, and adjustment in regard to work conditions and time). Every system providing values in the second tab are editable if a user also wants the result of estimation to be exported as a spreadsheet file (Fig. 5). This exported spreadsheet file shows the detailed computing procedure and estimation results as they appeared in the system. In addition, another sheet is automatically created in the exported file to show the information of the selected cases that are not considered in the retrieving step (Fig. 6).

Validation

Experiment Design

The MilFaCE system has simple user interfaces and the unique feature of user heuristic k (number of cases) selection for case adaptation. In this respect, the validation experiment is designed as follows: (1) to evaluate the system in terms of ease of operation and the accuracy of estimation results, we use nine facility types—quarters, mess halls, bathhouses, warehouses, magazines/arsenals, maintenance shops, office buildings and bachelor officers' quarters—that have at least 10 cases and extract the test data randomly using the random function of Microsoft *Excel*(Table 5); (2) at the same time, we arrange 10 military construction cost estimators, letting them estimate the cost of the selected cases using MilFaCE; and (3) finally, we question those estimators about the system's ease of operation and analyze the accuracy of the results.

The accuracy of the system-based estimation is compared to the traditional square-meter approach, which utilizes the average square-meter price and automatic five-NN based-methods. All cost values are converted to 2010 values. The absolute error ratio (AER) is defined in Eq. (4) (C_A and C_E = actual cost and estimated cost) to evaluate the effectiveness of the system and is compared to other counterparts

$$\mathrm{AER}(\%) = \begin{cases} \text{if } C_A - C_E > 1, & \text{then } [(C_A - C_E) - 1] \times 100 \\ \text{otherwise,} & [1 - (C_A - C_E)] \times 100 \end{cases} \tag{4}$$

Results and Discussion

As shown in Table 6, the system-based estimation resulted in an overall lower absolute error ratio than that of the traditional square-foot method. The deviations between the two test cases of the same facilities using MilFaCE were also lower than those

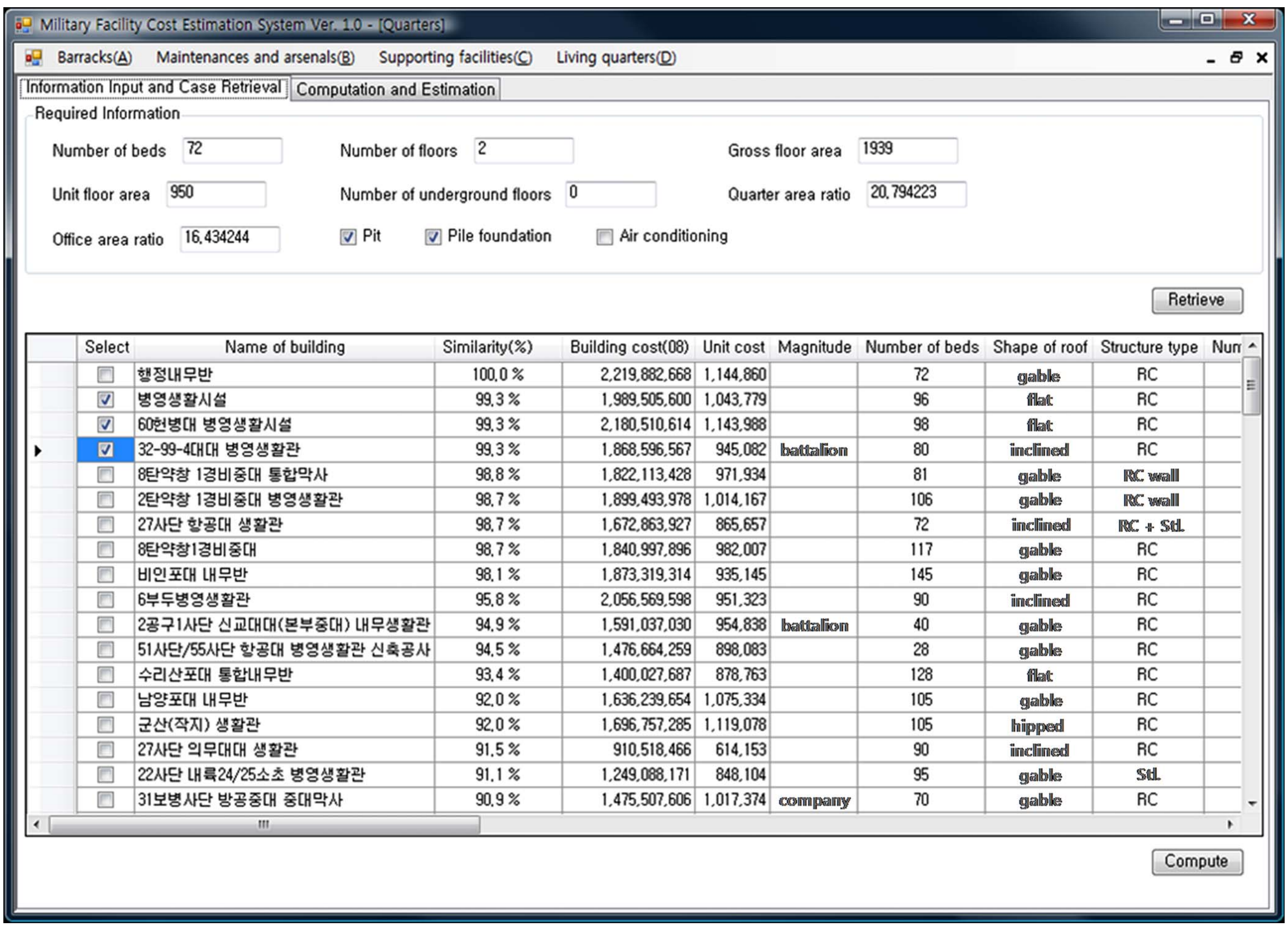


Fig. 3. User interface (information input and case retrieval tab)

J. Comput. Civ. Eng. 2011.25:218-231.

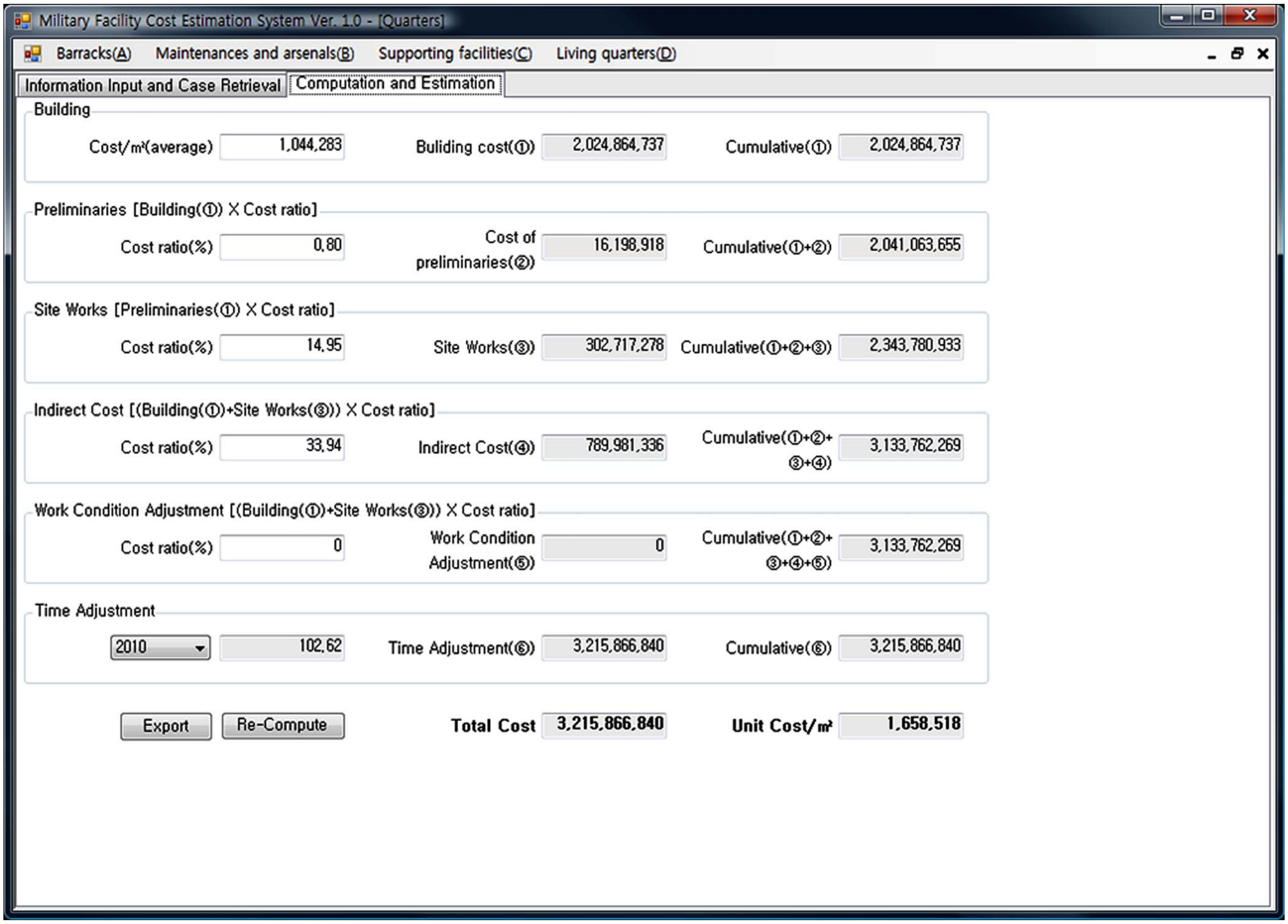
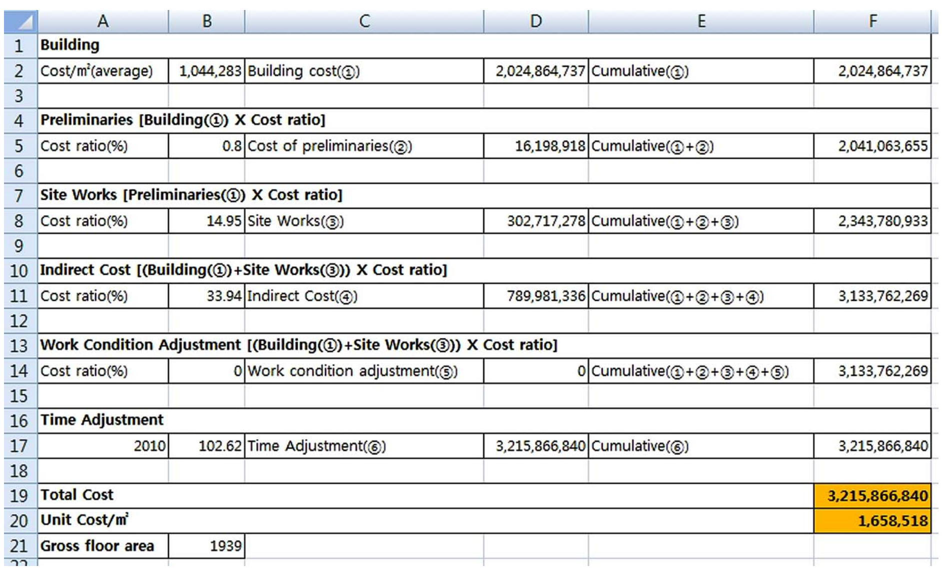


Fig. 4. User interface (computation and estimation report tab)



	A	B	C	D	E	F
1	**Building**					
2	Cost/m²(average)	1,044,283	Building cost(①)	2,024,864,737	Cumulative(①)	2,024,864,737
3						
4	**Preliminaries [Building(①) X Cost ratio]**					
5	Cost ratio(%)	0.8	Cost of preliminaries(②)	16,198,918	Cumulative(①+②)	2,041,063,655
6						
7	**Site Works [Preliminaries(①) X Cost ratio]**					
8	Cost ratio(%)	14.95	Site Works(③)	302,717,278	Cumulative(①+②+③)	2,343,780,933
9						
10	**Indirect Cost [(Building(①)+Site Works(③)) X Cost ratio]**					
11	Cost ratio(%)	33.94	Indirect Cost(④)	789,981,336	Cumulative(①+②+③+④)	3,133,762,269
12						
13	**Work Condition Adjustment [(Building(①)+Site Works(③)) X Cost ratio]**					
14	Cost ratio(%)	0	Work condition adjustment(⑤)	0	Cumulative(①+②+③+④+⑤)	3,133,762,269
15						
16	**Time Adjustment**					
17	2010	102.62	Time Adjustment(⑥)	3,215,866,840	Cumulative(⑥)	3,215,866,840
18						
19	**Total Cost**					**3,215,866,840**
20	**Unit Cost/m²**					**1,658,518**
21	**Gross floor area**	1939				

Fig. 5. Example of spreadsheet file output (computation and estimation report)

of counterparts. Interestingly, there is no relationship between experience years and the average AER of 18 test cases when they use the system. In other words, average AER does not get lower even though experience years increase. Estimation using 5-NN would be expected to yield a high degree of reliability only when many cases are available, and buildings are standardized, such as quarters, mess halls, and bachelor officers' quarters; otherwise, relatively high deviations result. On closer examination, it appears that when a large number of cases are available for certain facility types such as quarters (151), mess halls (35), and bachelor officers' quarters (50), estimate errors remain low regardless of the experience years of users. As shown in Table 6, case quarter-1 (CASE Q-1) and

	A	B	C	D
1	Name of building	병영생활시설	60헌병대 병영생활시설	32-99-4대대 병영생활관
2	Similarity(%)	0.993283879	0.993198117	0.992764665
3	Building cost(08)	1989505600	2180510614	1868596567
4	Unit cost	1043779.105	1143988.444	945081.6653
5	Magnitude			battalion
6	Number of beds	96	98	80
7	Shape of roof	flat	flat	inclined
8	Structure type	RC	RC	RC
9	Number of floors	2	2	2
10	Envelope materials	adobe	adobe	adobe
11	Gross floor area	1906.06	1906.06	1977.18
12	Unit floor area	1059.75	1059.75	988.59
13	Quarter area ratio	0.3173	0.317303757	0.5
14	Number of underground floors	0	0	0
15	Pit	TRUE	TRUE	TRUE
16	Office area ratio	0.1187	0.19264871	0.5
17	Dining area ratio	0.1148	0.128170152	0
18	Bathhouse area ratio	0.0378	0	0
19	Pile foundation	FALSE	FALSE	FALSE
20	Type of heating	hot water boiler	hot water boiler	hot water boiler
21	Air conditioning	TRUE	FALSE	FALSE
22	Year	2008	2008	2007
23	army/navy/air-force	army	army	army
24	Etc			

Fig. 6. Example of spreadsheet file output (detailed information of retrieved case)

quarter-2 (Q-2) have an average AER of 3.86% and 8.52%. For case mess hall-1 (CASE M-1) and mess hall-2 (M-2), the average AER is 7.46% and 10.01%, and for case bachelor officers' quarters-1 (CASE BOQ-1) and bachelor officers' quarters-2 (BOQ-2), 4.83% and 3.59%, respectively. In the case of warehouses, the AER is relatively higher than other building types. This is because, as revealed in Table 5, the selected test cases are either too large or too small in terms of size for consideration in the aggregated data range of warehouses. Hence, there are very few similar cases available, and this eventually leads to a limited selection of cases by users. Also as seen in the second case of warehouses (CASE W-2) the size of which is too small, despite small numerical differences between the estimation and the actual cost, the percentage difference appears to be huge.

With respect to the use of the buildings, in the case of quarters and bachelor officers' quarters which have relatively simple uses and standardized residential functions, the accuracy of estimates is high regardless of the estimators' experiences or career. Moreover, for some types of facilities such as mess halls, magazines/arsenals, maintenance shops, and office buildings that differ according to military conditions or functions, average costs estimated by MilFaCE are more accurate than those estimated by other methods. Regarding characteristics of a building (i.e., GFA), when the GFA is small and very few cases are available, estimation using the square-foot method has the lowest AER (i.e., car shade, bathhouses, and small-size warehouses), whereas the highest AER was obtained for magazines/arsenals and maintenance shops. Such results are likely to occur if the collected data have a biased distribution to a certain range of data sets. On the other hand, the system-based estimation has more stable deviations.

From the viewpoint of system convenience, all interviewees who have used MilFaCE responded that the system itself is very convenient for those with limited computer skills and experience. Furthermore, the estimation procedure is straightforward, and the system has a user-friendly interface. At the same time, they were satisfied with the estimation outcomes which were more accurate than the outcomes of traditional method. In addition, they placed a high value on the functions of the system that can be upgraded by simply uploading data and on the ability to customize their own data. Additionally, the interviewees responded that estimation based on the standard of 2010 leads to an insufficient budget compared to the actual financing required for military facilities.

Conclusions

Cost estimation at the initial stage of construction projects is crucial. Typically, estimations are performed repetitively in response to scope changes. Generally, those estimations are required to be reported in a brief time using only limited project information. In practice, the parametric method (e.g., square-foot method) is used most often. However, there is much effort involved in periodically (e.g., quarterly or annually) determining new unit prices, and those unit prices are often derived using data which do not consider the uniqueness of each project.

Consequently, there are limitations in terms of accuracy and explanation. The Korea Ministry of National Defense is no exception; there is great deal of difference between the parametric method estimations and actual expenditures. To deal with the conventional shortage in the Korean defense budget for construction projects, this research developed the MilFaCE system based on the CBR method, using case data from 422 military construction projects at 16 military facilities. The cost model has three components: (1) the CBR module for building cost; (2) the supplement module for site work, preliminary work, and indirect cost; and (3) the adjustment module for work conditions and escalation to the target year. Thereafter, we validated the effectiveness of the system in terms of the estimate accuracy and user friendliness.

The MilFaCE system, based on the cost model, is interlinked with the database. Hence, it is easy to update and customize. Overall, lower absolute error ratios (AER) were obtained by the system compared to the conventional square-foot method. It was also found that the system can help users who do not have sufficient knowledge of or experience in construction to make better decisions. The estimation procedure is straightforward, and the interface is user-friendly.

J. Comput. Civ. Eng. 2011.25:218-231.

Table 5. Profile of Cases for System Test

Facilities	Cost	ACC or NB or HH	NF	GFA	UFA	NUF	Pit	PF	AC	QR	OR	DR	HW	WR	OPR	RA	BH
Quarters																	
CASE Q-1	1,584,462,407	80	1	978	978	0.00	False	False	False	50.06%	11.50%						
CASE Q-2	7,069,829,942	260	3	5,434	2,214	0.00	False	False	False	26.20%	12.30%						
Mess halls																	
CASE M-1	803,739,690	168	1.00	551.62	559.30		False	False				54.20%					
CASE M-2	1,229,657,418	540	1.00	818.17	854.64		False	False				59.00%					
Bathhouses																	
CASE B-1	146,422,771			51.12													
CASE B-2	230,243,063			90.00													
Warehouses																	
CASE W-1	4,469,469,367		1.00	4,250.00	4,250.00		False	True	False							True	13.51
CASE W-2	55,652,185		1.00	40.00	40.00		False	False	False							True	4.15
Magazine and arsenals																	
CASE MA-1	61,685,166			22.08	22.08			False								False	3.00
CASE MA-2	85,143,253			91.00	91.00			False								False	3.20
Maintenance shops																	
CASE MS-1	907,858,421		1.00	540.00	540.00	0.00	False	False					10.54	53.33%			
CASE MS-2	1,530,481,175		1.00	910.34	910.34	0	False	False					18.25	55.24%			
Car shades																	
CASE C-1	120,460,958			252													
CASE C-2	167,784,907			351													
Office buildings																	
CASE O-1	1,061,118,679		2.00	622.32	314.88	0.00	False	False							0.00%		
CASE O-2	2,869,048,615		2.00	1,797.00	902.00	0.00	True	False							0.00%		
Bachelor officers' houses																	
CASE BOQ-1	325,658,262	6	1.00	173.70	175.24	0.00	False	False									
CASE BOQ-2	4,485,048,233	100	4.00	2,923.56	772.20	0.00	True	False									

Note: ACC = accommodation capacity; NB = number of beds; NH = number of households; NF = number of floors; GFA = gross floor area; UFA = unit floor area; NUF = number of underground floors; PF = pile foundation of building; AC = air conditioning; QR = quarter area ratio; OR = office area ratio to GFA; DR = dining area ratio to GFA; HW = height of work area of building; OPR = other purpose area ratio to GFA; RA = racks in a building; BH = building height.

Table 6. Comparison of Absolute Error Ratio (AER)

			Years of interviewee's career												
Facilities	Number of cases in DB	Name of test case	2	8	9	12	13	14	18	23	23	27	Average	cost/m^2	5-NN
Quarters	151	CASE Q-1	7.40%	0.69%	3.39%	1.43%	3.21%	4.53%	4.25%	1.86%	11.69%	0.18%	3.86%	2.44%	1.32%
		CASE Q-2	7.78%	10.11%	14.25%	8.86%	2.68%	1.06%	16.28%	13.36%	7.78%	3.05%	8.52%	8.47%	11.22%
Mess halls	35	CASE M-1	12.01%	1.91%	22.40%	3.81%	2.18%	6.15%	19.47%	3.13%	4.87%	0.51%	7.64%	12.38%	0.77%
		CASE M-2	2.94%	3.20%	11.25%	11.44%	3.52%	8.43%	4.11%	15.76%	14.48%	24.99%	10.01%	14.62%	9.39%
Bathhouses	19	CASE B-1	2.55%	3.05%	2.83%	9.96%	2.04%	7.43%	29.26%	2.83%	2.13%	24.04%	8.61%	12.61%	0.28%
		CASE B-2	8.67%	12.60%	9.18%	53.78%	9.81%	13.75%	1.23%	0.03%	20.41%	4.89%	13.43%	6.91%	0.27%
Warehouses	53	CASE W-1	30.89%	33.04%	38.82%	30.07%	26.54%	30.89%	10.01%	32.52%	32.52%	32.52%	29.78%	31.86%	32.65%
		CASE W-2	15.62%	45.58%	15.37%	30.18%	6.99%	18.11%	67.08%	21.43%	21.43%	67.43%	30.92%	9.07%	46.43%
Magazine/arsenals	26	CASE MA-1	25.58%	17.76%	37.59%	8.27%	1.21%	53.86%	65.37%	37.59%	6.38%	6.38%	26.00%	55.66%	26.65%
		CASE MA-2	4.27%	4.89%	4.36%	34.22%	5.22%	2.76%	5.09%	5.47%	5.09%	19.60%	9.10%	53.13%	0.05%
Maintenance shops	15	CASE MS-1	2.23%	8.93%	9.52%	8.93%	8.93%	11.16%	11.39%	13.40%	28.07%	8.93%	11.15%	53.92%	14.43%
		CASE MS-2	21.70%	36.74%	23.05%	22.14%	36.74%	26.85%	7.06%	43.42%	7.10%	36.74%	26.15%	42.16%	39.97%
Car shades	11	CASE C-1	0.83%	2.31%	20.45%	15.27%	0.85%	13.79%	10.31%	2.88%	6.94%	11.19%	8.48%	5.08%	0.67%
		CASE C-2	21.39%	6.25%	44.14%	25.69%	12.80%	28.47%	3.36%	49.69%	38.59%	6.85%	23.72%	1.43%	37.52%
Office buildings	24	CASE O-1	2.07%	5.90%	8.45%	2.07%	2.96%	5.08%	21.74%	8.82%	15.74%	27.49%	10.03%	11.52%	30.48%
		CASE O-2	9.21%	9.57%	14.30%	2.77%	9.57%	12.12%	9.57%	9.57%	9.57%	23.55%	10.98%	23.23%	2.81%
Bachelor officers' quarters	50	CASE BOQ-1	2.39%	1.25%	12.86%	3.95%	0.30%	9.17%	3.12%	3.73%	7.93%	3.57%	4.83%	0.83%	6.48%
		CASE BOQ-2	0.92%	0.12%	6.66%	0.92%	0.10%	1.34%	1.34%	0.12%	1.13%	23.23%	3.59%	19.84%	1.65%
Average			9.91%	11.33%	16.60%	15.21%	7.54%	14.16%	16.11%	14.75%	13.44%	18.06%	13.71%	20.29%	14.61%

J. Comput. Civ. Eng. 2011.25:218-231.

Consequently, the system can react more quickly and sensitively than traditional methods to changes in project information. The suggested CBR cost model makes the process of cost estimation more systematized and enhances the cost-planning method so that all participants can arrive jointly at a practical and efficient solution for the project, keeping within the budget. This enables a firm basis for decision-making in budget compilation. Also, the Korea Ministry of National Defense can plan and use their budget more efficiently by improving the budget standard. Actually, in late 2009, the Ministry had revised the sections related to construction budgeting in the Standard and Guide of National Defense Budget Organization of Korea. They adopted the process of the suggested cost model as the budgeting process and MilFaCE as the standard budgeting system for planning of military construction projects, and they revised the unit prices for facilities based on the analysis from the implemented database. Furthermore, this research presented the possibility for a commercial CBR cost-estimation system and would be a fundamental model for process- or methodology-building for CBR system development. Additionally, to benefit from the functions of the system, historical data must be continuously updated to improve the accuracy of estimates and to support decision-making for unique projects.

Because of issues of military security, the system was developed to be a stand-alone system. To ensure convenience and accessibility for a general purpose, the web-based system should be further developed. In addition, in-depth research is needed to determine the appropriate numbers of cases and to whom (what level of experience of users) it might be beneficial. Furthermore, it must be noted that this research is based on limited data from historical projects; therefore, the matter of bias toward the collected cases and outliers or noise-mixing in the aggregated data also must be taken into consideration.

Acknowledgments

This research was supported by grants (R&D06CIT-A03 and 05CIT-01) from the Korea Ministry of Land, Transport, and Marine Affairs and the Ministry of Defense.

Appendix.

Glossary of acronyms.

Acronyms	Full name	Acronyms	Full name
ACC	accommodation capacity	ICArchitecture	indirect cost, architecture
AC	air conditioning	ICCivil	indirect cost, civil
AER	absolute error ratio	ICMachinery	indirect cost, machinery
AT	area type of a unit household (e.g. 59 m^2, 84 m^2, and 114 m^2)	ICElectronic	indirect cost, electronic
BR	bathhouse area ratio to GFA	ICCommunication	indirect cost, communication
BCArchitecture	building cost, architecture	IR	interview area ratio to GFA
BCCivil	building cost, civil	LR	lodging area ratio to GFA
BCMachineryInside	building cost, machinery inside	NB	number of beds in a quarter building
BCMachineryOutside	building cost, machinery outside	NF	number of floors
BCElectronicsInside	building cost, electronics inside	NH	number of households
BCElectronicsOutside	building cost, electronics outside	NN	nearest neighbor
BCCommunicationsInside	building cost, communications inside	NUF	number of underground floors
BCCommunicationsOutside	building cost, communications outside	NUH	number of unit floor households
BH	building height	OR	office area ratio to GFA
BI	building information	OPR	other purpose area ratio to GFA
CBR	case-based reasoning	PF	pile foundation of a building
CI	cost information	PP	purpose of a building
CPWArchitecture	cost of preliminary works, architecture	QR	quarter area ratio to GFA
CPWElectronics	cost of preliminary works, electronics	RA	racks in a building
CPWMachinery	cost of preliminary works, machinery	SC	seating capacity
DR	dining area ratio to GFA	SI	site work information
EM	envelope material type	SR	shape of roof
GA	genetic algorithm	ST	structure type
GFA	gross floor area	SR	swimming area ratio to GFA
GI	project general information	TH	type of heating
HT	hallway type of a building (e.g. corridor or hall)	UFA	unit floor area
HW	height of work area of a building	WR	work area ratio to GFA

Note: HT (hallway type of building)—two types of hallways in Korea generally: corridor and hall; PP (purpose of building)—explains building use; three types of building use: warehouse, office building, and interview house; purposes of the warehouse can be hard goods, necessities, supplies, infectious waste, gas storehouse, and others.

J. Comput. Civ. Eng. 2011.25:218-231.

References

AACE Int. (1999). *Skills and knowledge of cost engineering*, 4th Ed., AACE, Morgantown, WV

Ahn, H., Kim, K., and HanI. (2006). "Global optimization of feature weights and the number of neighbors that combine in a case-based reasoning system." *Expert Syst.*, 23(5), 290–301.

An, S.-H., Kim, G.-H., and Kang, K.-I. (2007). "A case-based reasoning cost estimating model using experience by analytic hierarchy process." *Build. Environ.*, 42(7), 2573–2579.

Arditi, D., and Tokdemir, O. B. (1999). "Using case-based reasoning to predict the outcome of construction litigation." *Comput.-Aided Civ. Infrastruct. Eng.*, 14(6), 385–393.

Bode, J. (1998). "Neural networks for cost estimation." *Cost Eng.*, 40(1), 25–30.

Burkhard, H.-D. (2001). "Similarity and distance in case-based reasoning." *Fundamenta Informaticae*, 47(3–4), 201–215.

Chou, J.-S. (2009). "Web-based CBR system applied to early cost budgeting for pavement maintenance project." *Expert Syst. Appl.*, 36(2), 2947–2960.

Chua, D. K. H., Li, D. Z., and Chan, W.T.et al. (2001). "Case-based reasoning approach in bid decision-making." *J. Constr. Eng. Manage.*, 127(1), 35–45.

Chua, D. K. H., and Loh, P. K. (2006). "CB-contract: Case-based reasoning approach to construction contract strategy formulation." *J. Comput. Civ. Eng.*, 20(5), 339–350.

Creese, R., and Li, L. (1995). "Cost estimation of timber bridges using neural networks." *Cost Eng.*, 37(5), 17–22.

Goldberg, D. E. (2006). *Genetic algorithms in search, optimization, and machine learning*, Addison-Wesley, Boston.

Graza, J., and Rouhana, K. (1995). "Neural network versus parameter based application." *Cost Eng.*, 37(2), 14–18.

Hegazy, T., and Ayed, A. (1998). "Neural network model for parametric cost estimating of highway projects." *J. Constr. Eng. Manage.*, 124(3), 210–218.

Hoberman, S. (2005). *Data modeling made simple*, Technics Publications, Bradley Beach, NJ.

Jarde, A., and Alkass, S. (2007). "Computer-integrated system for estimating the costs of building project." *J. Constr. Eng. Manage.*, 13(4), 205–223.

Ji, S.-H. Park, M., and Lee, H.-S.(2009). "Cost estimation model for building projects using case-based reasoning. *Can. J. Civ. Eng.*, in press.

Karshenas, S. (2005). "Cost estimating in the age of 3D CAD software and object database", *Proc., Construction Research Congress 2005*, ASCE, Reston, VA.

Karshenas, S., and Tse, J. (2002). "A case-based reasoning approach to construction cost estimating." Proc., Int. Workshop on Information Technology in Civil Eng. 2002, *Computing In Civil Engineering*, 113–123.

Kirkham, R. J. (2007). *Ferry and Brandon's Cost Planning of Buildings*, 8th Ed., Wiley-Blackwell, Hoboken, NJ.

Kolodner, J. (1993). *Case-based reasoning*, Morgan Kaufmann Publishers, San Francisco.

Korea Ministry of National Defense. (2008). "Standard and guide of National Defense Budget Organization of Korea." ⟨http://www.mnd.go.kr/mndEng_2009/main/index.jsp⟩ (Nov. 1, 2009) (in Korean).

Lowe, D. J., Emsley, M. W., and Harding, A. (2006). "Predicting construction cost using multiple regression techniques." *J. Constr. Eng. Manage.*, 132(7), 750–758.

Mitchell, T. M. (1997). *Machine learning*, McGraw-Hill, Columbus, OH.

Morcous, G., Rivard, H., and Hanna, A., M. (2002). "Case-based reasoning system for modeling infrastructure deterioration." *J. Comput. Civ. Eng.*, 16(2), 104–114.

Ozorhon, B., Dikmen, I., and Birgönül, M. T. (2006). "Case-based reasoning model for international market selection." *J. Constr. Eng. Manage.*, 132(9), 940–948.

Pal, S. K., and Shiu, S. C. K. (2004). *Foundations of soft case-based reasoning*, Wiley-Interscience, Hoboken, NJ.

Qian, Z., Gao, W. S., Wang, F., and Yan, Z. (2009). "A case-based approach to power transformer fault diagnosis using dissolved gas analysis data." *Eur. Trans. Electr. Power*, 19(3), 518–530.

Ryu, H.-G. Lee, H.-S., and Park, M. (2007). "Construction planning method using case-based reasoning (CONPLA-CBR)." *J. Comput. Civ. Eng.*, 21(6), 410–422.

Sanders, G. L. (1995). "*Data modeling*, Course Technology, Inc., Cambridge, MA.

Smith, A. E., and Mason, A. K. (1996). "Cost estimation predictive modeling: Regression versus neural network." *Eng. Econ.*, 42(2), 137–161.

Sonmez, R. (2008). "Parametric range estimating of building costs using regression models and bootstrap." *J. Constr. Eng. Manage.*, 134(12), 1011–1016.

Soutos, M., and Lowe, D. J. (2005). "Procost—towards a powerful early stage cost estimating tool." Proc., 2005 ASCE Int. Conf. on Computing in Civil Engineering, *Computing in Civil Engineering*, ASCE, Reston, VA.

Tah, J. H. M., Carr, V., and Howes, R. (1998). "An application of case-based reasoning to the planning of highway bridge construction." *Eng., Constr., Archit. Manage.*, 5(4), 327–338.

Watson, I. (1997). *Applying case-based reasoning: Techniques for enterprise system*, Morgan Kaufmann Publishers, San Francisco.

Yau, N.-J., and Yang, J.-B. (1998). "Case-based reasoning in construction management." *Comput.-Aided Civ. Infrastruct. Eng.*, 13(2), 143–150.

Yi, J. et al. (2004). "A study on case-based forecasting model for monthly expenditures of residential building project." *Korean Journal of Construction Engineering and Management*, 79(1), 128–137.

J. Comput. Civ. Eng. 2011.25:218-231.