ARCHITECTURAL RESEARCH, Vol. 8, NO. 1(June 2006), pp. 77-84

The Application of CBR for Improving Forecasting Performance of Periodic Expenditures

– Focused on Analysis of Expenditure Progress Curves –

June Seong Yi
Department of Architecture, Ewha Womans University, Seoul, Korea

Abstract

In spite of enormous increase in data generation, its practical usage in the construction sector has not been prevalent enough compared to those of other industries. The author would explore the obstacles against efficient data application in the arena of expenditure forecasting, and suggest a forecasting method by applying Case-based Reasoning (CBR). The newly suggested method in the research, enables project managers to forecast monthly expenditures with less time and effort by retrieving and referring only projects of a similar nature, while filtering out irrelevant cases included in database. Among 99 projects collected, the cost data from 88 projects were processed to establish a new forecasting model. The remaining 10 projects were utilized for the validation of the model. From the comprehensive study, the choice of the numbers of referring projects was investigated in detail. It is concluded that selecting similar projects at 12~19 % out of the whole database will produce a more precise forecasting.

The new forecasting model, which suggests the predicted values based on previous projects, is more than just a forecasting methodology; it provides a bridge that enables current data collection techniques to be used within the context of the accumulated information. This will eventually help all the participants in the construction industry to build up the knowledge derived from invaluable experience.

Keywords: Expenditure, Forecasting Model, Cost Data, CBR (Case Based Reasoning), Apartment Housing, Computerization

1. INTRODUCTION

Accurate information about the expenditures enables the managers to capture the problems as early as possible and to correct them with less difficulty. Along with planning incomes, forecasting expenditures is indispensable for ensuring a high level of credit reliability for a company. A cash flow statement indicates the movement of cash-in (income) and cash-out (expenditure) of a company over a certain period of time. The accuracy of the forecast depends on the precision of the prediction of the periodical income and expenditure.

As market competition becomes more intense, the importance of a comprehensive assessment of financial status has become crucial for both the success of the project and the survival of the company. Despite its importance, cash flow planning, which is a part of the financial management, has not received enough attention from the construction industry. The reason for this is the ever-changing environment of construction projects, and consequently, the large amount of work involved in compilation and updating of such a plan.

With the application of statistical models, databases, and machine learning, researchers began to develop a set of new tools and techniques to analyze very large databases. However, most computer programs have been designed for processing the current independent project. Data from historic projects are invaluable asset for all the participants involved. If all this information were systematically structured to allow easy access to planners, it would enhance the certainties, as projects progress. In many cases, there is little or no useful historical information available at the time the initial forecast is required. Thus, the early forecast must be based largely on subjective considerations. As more information becomes available, managers must modify their subjective estimates in terms of the actual data. However, the existing model mostly depends only on historic data without considering the current trends. The need for measuring the change over time for the current project is demanded in a rapidly fluctuating situation.

Most of the experienced knowledge extracted from a project, becomes useless without the interpretation by those people who were directly involved in the project, because they are the only ones who experienced and controlled all the tasks during the process. The obstacles in developing an appropriate model, which represents the nature of the construction industry, are derived from the following reasons: 1) a project-based system in the construction industry prevents sharing information and learning from historic data, 2) the lack of a standardized tool for analyzing the project progress exacerbates the efficiency in inheriting knowledge between projects, and 3) ignorance of the work-package structure as a project control tool misleads managers in capturing the problem.

Existing forecast models have weaknesses in capturing the significant variance derived from the set-off between elements, and work-packages. The Case-based reasoning (CBR) is regarded as the proper tool to resolve the problems listed above. It uses the knowledge represented in specific cases to solve a new problem based on the similarities between the new problem and the available cases.

The purpose of this research is to develop the dynamic forecasting model for the periodic expenditure over construction periods, which provides the basic information for the financial status of the project by applying Case-based reasoning (CBR). To overcome the above-mentioned deficiencies and help managers make appropriate decision, the following factors are taken into consideration: 1) forecast based on information derived from respective work-packages, 2) diagnosis of sources of excessive variances, 3) generation of a standardized distribution function for each work-package, 4) providing the standardized s-curve for monthly expenditure throughout all construction periods, and 5) suggestion of the number of selected projects among the whole cases in the database for more accurate expenditure prediction. Another objective is to extract as much knowledge as possible from similar projects, which saves significant time and effort and increase the efficiency in managing projects.

2. LITERATURE REVIEW

Comprehensive reviews on the previous studies were addressed regarding cost forecasting models mainly focused on two subjects; cash flow, case-based reasoning

2.1 Cash flow

The importance of cash flow forecasting cannot be overemphasized. Cash flow forecasting provides valuable early warning systems which enable managers to recognize problems and solve them. In calculating cash flow, a variety of information, which includes a number of items, resources, time intervals, payment periods, payment delay, amounts of retention, and the project cost breakdown structure, is required. The calculation process demands tedious tasks, which necessitated the use of computers. Peterman (1972) developed a computer model to forecast contractors' cash flow based on contract bar charts and unit price information. Allsop (1980) linked the unit costs from estimating program to obtain cash flow. Ashley and Teicholz (1977) developed a model base on the value curve to assist in the analysis of cash flow throughout the whole project life. Oliver (1984) analyzed projects collected from three construction companies and concluded that projects are individually unique and pursued respective value curves based on historical data. Kenley and Wilson (1986) developed the value s-curve assuming that there should be errors caused by discrepancies from the average trend line rather than random error. Consequently, they claimed that each project has an individual line of central tendency. This concept was continuously explored and developed to claim that individual building projects with different sizes and types have their own characteristic cash flow. Peer (1982) used regression techniques to produce a cash flow forecasting prior to construction. Berny and Howes (1982) suggested the integrated model of polynomial and exponential curve for expenditure forecast. Singh and Woon (1984) exploited the features existing in the cash flow, and concluded that there is correlation between the project type and its expenditure pattern. Pattern (1982) emphasized the importance of cash flow information as a management tool. He presented a dynamic model for forecasting characteristics of a typical construction project. The model used the historical data as the basis of its information.

2.2 Case-based reasoning (CBR)

Case-based reasoning (CBR) is a major paradigm in automated reasoning and machine learning. In CBR, the system solves a new problem by noticing its similarity to one or several previously solved problems and by adapting their known solutions instead of working out a solution from scratch. The process to describe CBR can be explained as followings:

- Retrieve the most similar case(s).
- Reuse the case(s) to attempt to solve the problem.
- Revise the proposed solution if necessary
- Retain the new solution as a part of a new case.

According to these steps, Aamodt and Plaza (1994) describe a CBR as a cyclic process comprising "the 4 RE's"; Retrieve, Reuse, Revise and Retain.

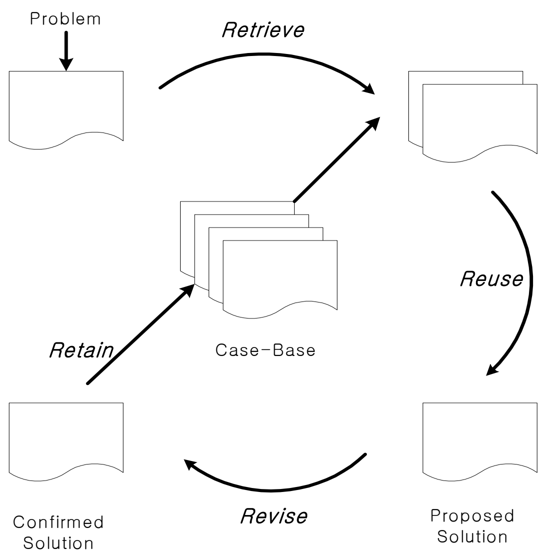

Figure 1. Case-based reasoning cycle

Stottler (1992) applies a CBR system for cost and sales prediction under uncertainty, which leads to an increase in business profit. He tried to retrieve the similar example in the past, and use that knowledge for business planning, e.g., staffing, sale volume, financing, etc. Lee et al. (1995)

describes a CBR system that develops forecasts for cash flow accounts. From the proposed system in his research, fuzzy integrals are used to calculate the synthetic evaluations of similarities between cases instead of the usual weighted mean.

3. DEVELOPMENT OF THE FORECASTING METHODOLOGY BASED-ON CASE-BASED REASONING

Construction managers are supposed to make decisions based on their experience and newly generated data as projects progress. In many cases, there is little or no useful historical information available at the time the initial forecast is required. Thus, the early forecast must be based largely on subjective considerations. As more information becomes available, managers must modify their subjective estimates in terms of the actual data. However, the existing model mostly depends only on historic data without considering the current trends. The need for measuring the change over time for the current project is demanded in a rapidly fluctuating situation.

The purpose of developing the new forecasting methodology is to generate the precise prediction of periodic expenditure as a project proceeds by choosing only similar projects as references instead of all cases in a database. The theoretical framework of the research is described in the following section.

3.1 Theoretical framework

The primary research concern will be focused on finding the efficient way of matching the similar projects in the past to the current one. The new forecasting method suggested in this research follows four steps: 1) data standardization of all the data from various projects; 2) data transformation to generate expenditure pattern curves for respective projects in the database to synchronize with the current project which would be forecasted; 3) matching the progress pattern of current project to those of the historic projects in the database to retrieve the similar projects ; 4) the development of an expenditure pattern curve for the current project in progress based on data derived from similar projects in the past.

The first step of the process is to processed and standardized monthly cost data into percentiles from monetary amounts. Thereafter, another conversion is performed to synchronize the different time frame into the same one with the current project. The second step is to develop the respective expenditure pattern curves of individual projects whose time scale is modified to coincide with the duration of the current project in progress. The case matching is applied in the third step for examining the similar projects compared with the current project in progress. The output of matching process is derived from the sum of squares of differences (SSD) between matched cases. At this stage, all the projects in the database are arrayed in the order that a project with smaller SSD comes first. The SSD value is the key concept in this research, which provides a guide for similarity evaluation. The definition and acquisition process for SSD is presented in the section 3.2 similarity evaluation. Finally, forecasted expenditure curves are generated in the fourth step. For its validation, the results of new methodology based on case-based reasoning (CBR) are compared with those of the traditional method that refers to the averaged values of all the projects in database. The appropriate number (or proportion) for selecting projects from the database will be determined at this stage.

3.2 Similarity evaluation

As previously mentioned, the success of the new model depends on the efficiency and properness of retrieving similar project(s) among the past projects stored in the database. There could be various criteria to define similarity of residential building projects, which are gross floor area, the number of housing units, the number of building units, monetary amounts, the number of storey, construction duration in months, and so on. In this study, the pattern of expenditure curve was explored to prove its significant role to characterize the project. It is assumed that thorough analysis on monthly expenditure focused on work-package level will lead to clear understanding of the project characteristics.

The ultimate objective of the new methodology is to predict monthly expenditure of the on-going project based on the cost data at the work-package level from past projects. Since most historic projects in the database are probably different in terms of monetary amounts and construction duration, it is necessary to develop a process to standardize the variables corresponding to those of the current project.

The first step of the process is to convert monthly cost data into percentiles from monetary amounts. Thereafter, another conversion is performed to synchronize the different time frame into the same one with the current project. The SSD could be obtained when all the historic data are transformed and standardized. The differences between the matched cells are squared, and respective value at each month will be added up to the point when the prediction is performed. When the SSD turns out ‘0’, two projects are completely identical from the view of expenditure pattern. From the same point of view, two projects with much smaller SSD presents the more similarity than other matches. The processes of data manipulating and calculating SSD are presented in figure 3 and formula (1) respectively.

3.3 Data processing

1) Characteristics of collected data set

The collected data to build the systems were limited to residential building projects with lump sum contracts and a design-bid-build delivery type. The research was also limited to projects larger than ₩10 billion. The summary of 88 projects collected for this research are shown in Table 1.

Some assumptions had to be made to maintain the problem at a manageable level. First, no cost indices compensating the difference in time frame were applied in comparing expenditure data between the projects in the data set. Though the researcher established the time laps between 1995 and 2003 as the horizon for this research, any modification based on the different fluctuations in material, wages, and equipment costs were not taken into account. All the values represented in terms of proportion against the costs at completion are able to standardize the data over the different time frame based on the assumption that the identical inflation rate were applied to the prices among the categories: e.g., material cost, wages, and equipment rental costs, etc.

Secondly, it was assumed that all monthly cost data would be obtained on the last day of the month for which they would be accumulated up to that point. Lastly, the possible difference derived from various construction areas was disregarded in this research. To obtain a more reliable result, the distinction according to the location - the urban area and suburban area - should be reflected in processing the cost data.

Table 1. Characteristics of projects in database

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Duration	88	29.386	29.000	29.288	4.290	0.457
Floor Area	88	907270	638573	790516	897288	95651
No. Base	88	1.6364	2.0000	1.6000	0.6640	0.0708
No. Floor	88	20.307	20.000	20.263	4.499	0.480
No. Bldg	88	7.091	6.000	6.538	5.364	0.572
No. Unit	88	591.9	452.5	563.9	401.6	42.8
Cost	88	41829	34154	39568	27557	2938

2) Data format

The data required for the model were the monthly actual cost as well as the project summary. The structure and elements of the raw data were based on those of the company that provided the data(see to Table 2). Data from 88 completed residential building projects were analyzed for the purpose of applying the new forecasting methodology.

Table 2. Data structure

7-major work package	20-work package
Building	Temporary, Foundation, Concrete, Masonry, Waterproof, Plaster, Tile/stone, Wood/carpentry, Window, Floor/wall finish, Paints, Furnishing, Sundries, Equipment
Electrical, Mechanical, Earthwork, Planting, Annex Building General Expenses	Electrical, Mechanical, Earthwork, Planting, Annex Building General Expenses

All the monthly-based cost data have been transformed into percentile value against the total cost to accomplish standardization. Continued data collection using the methodology developed will serve to enable a more precise forecasting in the future. The primary focus of this research is to predict monthly expenditure by referring to the similar expenditure patterns of past projects in database. To match the current project to the similar ones, it was necessary to develop a process to standardize the variables and project durations.

3) Generating the standard curve of expenditure pattern

A wide range of statistical techniques was utilized to predict dependent variables regarding expenditure in residential building projects. Establishing the cumulative curve as well as a progress distribution is very important. A significantly large sample size is required to obtain an accurate curve. Progress at each monitoring point is calculated and assembled into the standardized curve. The curve obtained from the sample might be rough, and some curves might be inconsistent with the overall curve shape. Nevertheless the discrepancy between the individual projects' S-curves and standardized one, the standardized S-curve provides a base for controlling and planning at early stage. Progress distribution over time and cumulative curve of each work package is developed based on averaged data from the 88 dataset. The figure 2 presents exemplary pattern of plaster work package, one of 20 work packages investigated in this study.

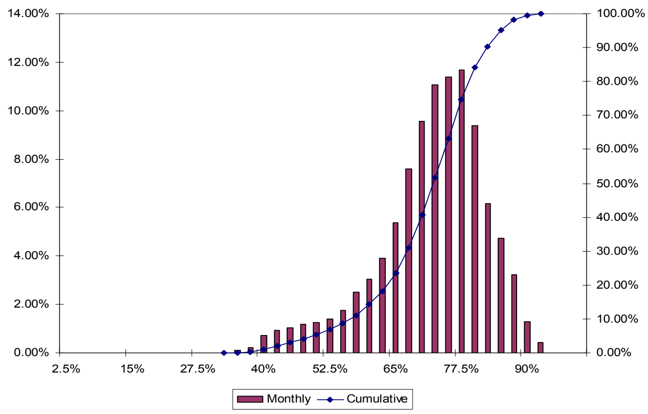

Figure 2. Periodic expenditure pattern and cumulative curve (plaster work package)

3.4 Model development using new methodology

In the development of the model, the first stage is to define the structure of data set. The cases here will represent individual projects of which components are respectively monthly expenditure patterns from the 20-work package level. The format of input data should hold the consistency complying with the current accounting system by which projects are controlled.

The data analysis tool consists of four modules: 1) data standardization; 2) data transformation; 3) quantitative analysis and matching similar cases; and 4) forecast generation. These four modules comprise the model that is developed in this research. More detailed descriptions for each model components are depicted below.

1) Data standardization

The first module, data standardization, is aimed at avoiding extreme modification from the existing system. It is designed to provide an easy interface by which an end-user can input data into the system. The first step of the program input request is to convert all the cost data of the past projects into percentile against the final account cost from monetary value. Next, the user is given the opportunity to designate the construction duration in months according to that of the current project

2) Data transformation

The data transformation module performs the transformation of the time variable for the horizontal axis and the percentage complete for the vertical axis. The need to generate the data of a consistent number of data points for each project is crucial for the success of the new forecasting model. The number of data points being generated is determined by the duration of the current project in progress. When the duration of the currently monitored project is 30 months, all the data points of previous projects in the database will be converted into 30 splits by applying either interpolation or extrapolation. From the expenditure curve for each work package, the assumption, that costs were expended in a linear fashion between the two adjacent points, is acceptable. The number of splits of all projects in the database is synchronized with the time frame of the project in progress.

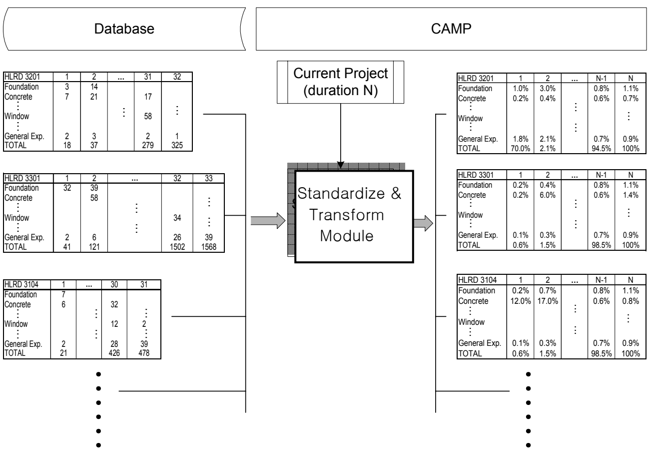

Figure 3. Example of data standardization and transformation (construction duration: N months)

3) Retrieving the similar cases by quantitative analysis

The main performance at this stage is to compare the current project to previous ones in the database, and retrieve similar cases for a more precise forecasting. The degree of similarity between the current projects and ones in the database is determined by the sum of squares of difference (SSD). The case-matching module reads the complete set of projects in the database and convert them into the identical time frame with the project in progress. This process provides a more accurate forecast by referring to only similar projects based on expenditure values at each work package level. The formula (1) and Figure 6 explain how to calculate SSD of the 20- work package level analysis. Afterward, a calculation process takes place, and finally forecasts the remaining time points in advance based on the values derived from the similar projects. The appropriate proportion for project selection as references will be further addressed in the model validation part.

$$SSD = \sum_{i=1}^{I}\sum_{j=1}^{n}(W_f ij - W_d ij)^2 \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

Where,

W_f : progress percentile of currently forecasted project
W_d : progress percentile of projects in database
i : the number of work package classified(i=1, 2, … , I), for 20- work package level analysis, the value of I will be 20.
j: month (j=1, 2, …, n),
n: the n^{th} month from the start, when forecast executed

Currently forecasted project W_f

		Month				
i	Description	1	2		n − 1	n
1	Temporary	$W_{f\,1,1}$	$W_{f\,1,2}$	……	$W_{f\,1,n-1}$	$W_{f\,1,n}$
2	Foundation	$W_{f\,2,1}$	$W_{f\,2,2}$	……	$W_{f\,2,n-1}$	$W_{f\,2,n}$
3	Concrete	$W_{f\,3,1}$	$W_{f\,3,2}$	……	$W_{f\,3,n-1}$	$W_{f\,3,n}$
⋮		⋮	⋮	……	⋮	⋮
15	Electrical	$W_{f\,15,1}$	$W_{f\,15,2}$	……	$W_{f\,15,n-1}$	$W_{f\,15,n}$
16	Mechanical	$W_{f\,16,1}$	$W_{f\,16,2}$	……	$W_{f\,16,n-1}$	$W_{f\,16,n}$
⋮		⋮	⋮	……	⋮	⋮
20	General Exp.	$W_{f\,20,1}$	$W_{f\,20,2}$	……	$W_{f\,20,n-1}$	$W_{f\,20,n}$

Projects in database W_d

		Month				
i	Description	1	2		n − 1	n
1	Temporary	$W_{d\,1,1}$	$W_{d\,1,2}$	……	$W_{d\,1,n-1}$	$W_{d\,1,n}$
2	Foundation	$W_{d\,2,1}$	$W_{d\,2,2}$	……	$W_{d\,2,n-1}$	$W_{d\,2,n}$
3	Concrete	$W_{d\,3,1}$	$W_{d\,3,2}$	……	$W_{d\,3,n-1}$	$W_{d\,3,n}$
⋮		⋮	⋮	……	⋮	⋮
15	Electrical	$W_{d\,15,1}$	$W_{d\,15,2}$	……	$W_{d\,15,n-1}$	$W_{d\,15,n}$
16	Mechanical	$W_{d\,16,1}$	$W_{d\,16,2}$	……	$W_{d\,16,n-1}$	$W_{d\,16,n}$
⋮		⋮	⋮	……	⋮	⋮
20	General Exp.	$W_{d\,20,1}$	$W_{d\,20,2}$	……	$W_{d\,20,n-1}$	$W_{d\,20,n}$

Figure 4. Example of formatted data for case matching (20- work package level; i=20)

4) Forecasting

The forecast generation module then calculates the average values derived from the selected projects, and fills out the remaining blank data points with the average values of those matched projects. The model validation module enables users to modify the number of referring projects by their own decisions. The quantity for the ordinate axis is the value as a percent of the total expenditure. The horizontal axis measures time in months from the beginning of the construction.

3.5 Computerization of the model

1) Considerations for computerization

It was necessary to perform numerous tasks for handling data to develop the new forecasting model. Executing all of the steps required to perform mathematical calculations and developing the pattern curves would have been extremely time consuming. The flexibility and efficiency in recreating the analysis as new data is essential to the success of the model. Microsoft Excel and Microsoft Visual Basic were chosen to satisfy the requirements that are mentioned above. These two packages play their roles in handling data input, producing the needed graphical output, and performing statistical analysis. Several other considerations for the computerization of the model are depicted as below.

- ease of use by end user;
- generating easy to understand numerical and graphical output;
- easily convertible to different types of projects;
- flexibility to add new data into the database.

2) Graphical User Interface (GUI)

The GUI of the model is devised to provide easy access for the users. A brief explanation for main component is depicted in figure 5.

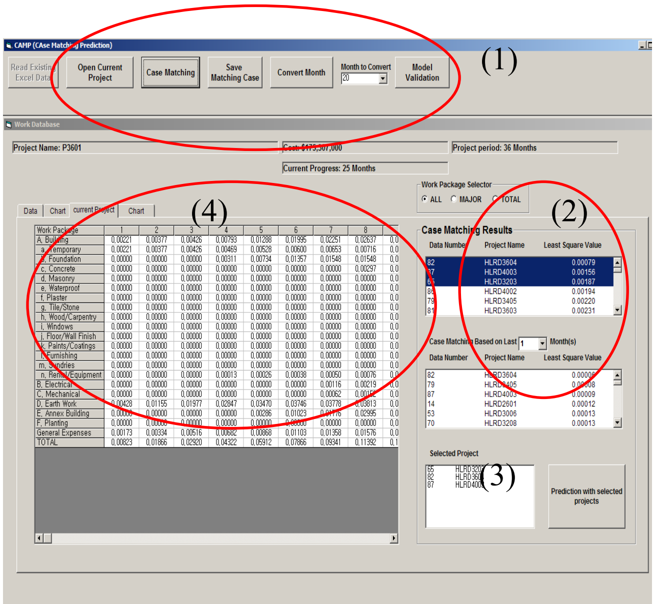

Figure 5. Exemplary GUI

- Components indicated by (1) perform data loading and manipulating at the beginning of forecasting task(see to section 3.4).
- Component (2) presents the list of past projects which are arranged by the result of SSD calculation(the project with smaller value is listed upper).
- Component (3) shows the selected projects as references in forecasting monthly expenditures of the current project. The decision process for the number(or percentage) of projects selection is explained in the following chapter 4.
- Component (4) is the final output, which is the forecasted monthly expenditure for the current project up to the point of expected completion date.

4. VALIDATION OF THE NEWLY PROPOSED METHODOLOGY

The purpose of this section is to show how well the newly developed forecasting methodology predicts values for monthly expenditure. The main objective at this phase is to figure out the appropriate number in selecting projects to obtain the forecasting reference. The above analyses are conducted at three different forecasting periods of times, which are 35%, 50% and 70% completion of the project. To prove the effectiveness and accuracy of the new methodology, the results of newly suggested method in this study will be compared with those of a conventional methodology that refers to the average of the whole database.

4.1 Measure of forecasting performance

10 completed projects other than database will be selected and evaluated for the purpose of model validation, when the proportion of case selection and the choice of work package level among three analyses (20-work package level, 7-major work package level, and total sum level) were decided. Two separate modules are required to evaluate the model performance by *% error* over the remaining construction duration and thereafter the forecasting point. One is to produce the forecast value at i^{th} month (X_{Fi}), which is averaged from the selected projects. It is calculated by employing the formula (2) below:

$$X_{Fi} = \frac{\sum_{j=1}^{n} X_{Fij}}{n} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Where,

n: the total number of selected projects as references
j: selected projects as references

The other module is to generate *% error* at the forecasting time, i^{th} month. It is calculated by employing the formula (3) below:

$$\%\ error\ at\ i^{th}\ month = \frac{|X_{Fi} - X_{Ai}|}{X_{Ai}} \times\ 100\ \ldots\ldots (3)$$

Where,

X_{Ai} : actual data at i^{th} month,
X_{Fi} : forecasted value at i^{th} month

Above two modules are integrated into one and provide overall model performance from the forecast conducted month *k* to the end of the construction period *N*. It is calculated by employing the formula (4) below:

$$\%\ error = \frac{\sum_{i=k}^{N} \frac{\left| \sum_{j=1}^{n} \frac{X_{Fij}}{n} - X_{Ai} \right|}{X_{Ai}}}{(N - k)} \times\ 100\ \ldots (4)$$

Where,

N: construction duration,
k: the k^{th} month from the start, when forecast executed

4.2 Evaluation on forecasting performance
1) Forecasting performance at different phases

The issue addressed in this section is to figure out the appropriate number(or proportion) of projects selection that provides the most accurate forecasting for monthly expenditure. The forecasting performances at three different phases (35%, 50%, and 70% completion of the projects) are measured and analyzed, as the number of projects selected increases from 1 up to 88. At each forecasting point, its analysis was respectively classified into three different levels, which include 20-work package level, 7-major work package level, and total sum level. The project in progress is matched to the similar projects according to the results of the sum of squares of difference (SSD).

The summary of *% error* results from testing the case project is given in Figure 6, as the number of projects selection increases 1 to 88. In its matching process, all the projects in database are ranked by the SSD values from 20-work package level. Regarding the forecasting evaluation from 35% completion, it can be seen that the *% error* decreases as the number of selected projects increases until the adequate model is reached. Therefore, the adequate model for this dataset is to choose 12 projects as base reference for forecasting residential building projects. When the user selects 12 similar projects out of an 88 project dataset, the newly developed model provides the prediction of monthly expenditure with a significant value of 0.76 *% error*. If all 88 projects are averaged as in traditional method, the value of 5.82 *% error* will be obtained.

The very same processes were repeated both for the phases of 50% and 70%. It is encouraging that both results at other forecasting points (50% and 70%) also represent significant improvements in forecasting the monthly expenditure. When prediction was conducted at the 50% completion time, the prediction accuracy was improved up to 0.69 *% error*, with the averaged curve of referenced 17 projects. The *% error* would be 4.21%, if all 88 projects were averaged in generating the standardized curve. Similarly, the accuracy for the forecast at 70% completion time was improved from 2.67 *% error* up to 0.78 *% error* by choosing 9 projects instead of 88 projects.

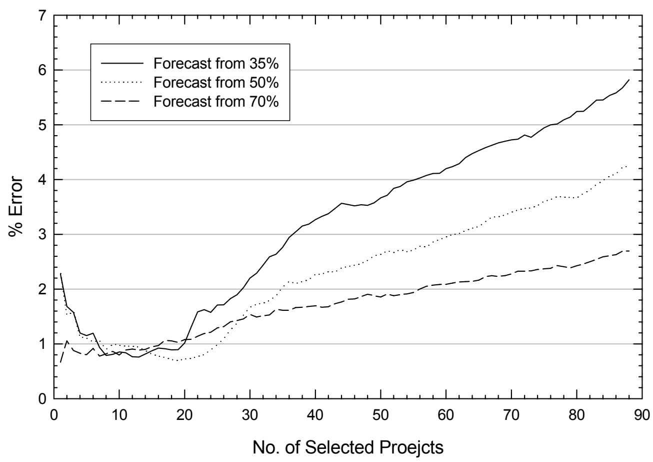

Figure 6. *% error* Summary at Three Different Phases
(20-work package level analysis)

5. CONCLUSION

This study has taken an in-depth and focused look at one central issue: enhancing the forecast accuracy for monthly expenditure of residential building project. This issue was addressed through the use of the Case-based reasoning (CBR) by matching and retrieving similar projects from past projects. Result of the analyses indicated that the forecasting performance could be significantly improved by applying the new model, which is developed in this study.

The developed model produces forecasting with an increase in accuracy of 0.76 *% error* at the point of 35 % progress, which would have been 5.82 % in the traditional methodology. Similar improvements are shown in from the point of 50 % and 70 % progress, which are 0.69 from 4.21 *% error*, and 0.78 from 2.67 *% error* respectively.

The primary contribution of this research is to increase the forecasting accuracy for expenditure pattern throughout the project. The application of the method suggested in this research is useful for forecasting the expenditure of residential building projects as they progress. Besides saving time, expenses, and the efforts of the project managers, there are several contributions both to the academic area and the construction industry. These include extending the forecasting horizon, applicability to different monitoring points, and enhancing the knowledge build-up and data exchange by referring to the similar projects in the past. The practical contributions are outlined below.

- extending the forecasting horizon
- deeper insights into the status of project
- generation of expenditure pattern curves of each work package
- efficiency improvements in knowledge build-up and data exchange

The new model, which suggests the predicted values based on similar projects in the past, has meaning beyond just a forecasting methodology. It provides a bridge that enables current data collection techniques to be used within the context of the accumulated information. This will eventually help all the participants in the construction industry to build up the knowledge derived from invaluable experience.

REFRENCES

Aamodt, A. and E. Plaza (1994). “Case-Based Reasoning: Foundational Issues, Methodological Variations and System Approaches.” AI Communications 7: 39-59.

Allsop, P. (1980). “Cash Flow and Resource Aggregation from Estimators’ Data.” Construction Management, Loughborough University of Technology.

Ashley, D.B. and P. Teicholz (1977). “Pre-Estimate Cash Flow Analysis.” Journal of Construction Division 10(3): 111-113.

Berny, J. and R. Howes (1982). “Project Management Control Using Real Time Budgeting and Forecasting Models.” Construction Papers 2: 19-40.

Kenley, R. and O. Wilson (1986). “A Construction Project Cash Flow Model – an Ideographic Approach.” Construction Management and Economics 4: 213-232.

Lee, R.W., Barcia, Ricardo Miranda, and Khator, Suresh K. (1995). “Case-Based Reasoning for Cash Flow Forecasting Using Fuzzy Retrieval.” First International Conference (ICCBR – 95).

O’Leary, T. and S.N. Tucker (1996). “Techniques and Tools for Project Cash Flow Prediction in the Australian Construction Industry.” Proceeding of 40th Annual Meeting of AACE, Vancouver, Canada.

Oliver, J.C. (1984). “Modeling Cash Flow Projections Using a Standard Micro Computer Spreadsheet Program.” Construction Management Projects, Loughborough University of Technology.

Pattern, W.N. (1982). “Application of Cost Flow Forecasting Models.” Journal of Construction Division 17128, 108 C02(3): 226-232.

Peterman, G.G. (1972). “Construction Company Financial Forecasting.” Arizona State University.

Singh, S. and P.W. Woon (1984). “Cash Flow Trends for High-rise Building Projects.” Proceedings of the 4th International Symposium on Organization and Management of Construction, University of Waterloo, Canada.

Stottler, R.H. (1992). “Case-Based Reasoning for Bid Preparation.” AI Expert(March): 44-49.

(Data of Submission : 2006. 2.10)