Automation in Construction 18 (2009) 386–393

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Artificial intelligence approaches to achieve strategic control over project cash flows

Min-Yuan Cheng [1], Hsing-Chih Tsai *,[2], Chih-Lung Liu [3]

Department of Construction Engineering, National Taiwan University of Science and Technology, Taiwan

ARTICLE INFO

Article history:
Accepted 17 October 2008

Keywords:
Construction industry
Project success
Cash flows
Artificial intelligence
K-means clustering
Genetic algorithm
Fuzzy logic
Neural network

ABSTRACT

The ability over the course of a construction project to make reliable predictions regarding cash flows enhances project cost management. This paper uses artificial intelligence (AI) approaches to predict cash flow trends for such projects in order to develop appropriate strategies that apply factors such as float, process execution time, construction rate and resource demand to project cash flow control. AI approaches involved in this paper include K-means clustering, genetic algorithm (GA), fuzzy logic (FL), and neural network (NN). K-means clustering is employed to categorize similar projects, while the other approaches are used to develop the Evolutionary Fuzzy Neural Inference Model (EFNIM), a knowledge learning model. FL and NN are employed in the EFNIM to develop a neural-fuzzy model that can deal with uncertainties and knowledge mapping. GA is used to optimize the membership functions of FL and NN parameters globally. The major target of this AI learning is to address sequential cash flow trends. This trained result is furthermore applied to a strategic project cash flow control. This cash flow control affects project performance within the banana envelope of the S-curve for project management.

1. Introduction

Cash is a critically important construction resource. Profitability can greatly influence project cash flow during project progress. Russell [32] has indicated that over 60% of construction contractor failures are attributable to economic factors. Cash flows control impacts significantly upon project management and project success. Cost flow management involves forecasting and control. Dynamic cash flow forecasts based on financial considerations provide the advanced knowledge necessary to manage cash flows. Various methods have been applied to cash flow management based on analyses of real business environments.

Almond and Remer [1] presented six continuous cash flow models with two special limiting cases. As forecasting cost flow at the company level is more difficult than at the project level, Navon [30] developed a system to generate company cash flows automatically. Chen et al. [6] recommended an extended cost-schedule integration model influenced by more detailed payment conditions, including differential payment lags, components for materials and labor, and payment frequency. Park et al. [31] adopted moving weights of cost categories to build a realistic cash-out model and transferred net planned monthly earned values as cash-in simply to forecast cash flow that reflects time-lagged impacts during a construction phase. Khosrowshahi and Kaka [21] developed a model that integrates modules of an exponent and two fourth-degree polynomials to forecast and manage project cash flow. Gormley and Meade [12] proposed a time series model to provide expectations of further cash flow. Jarrah et al. [18] used a fourth degree polynomial regression analysis for TxDoT highway project payment curve forecasting to predict cash flows and trends.

Most project cash flow models use S curves to forecast project cash flows because S curve application is the simplest method to project cash flows [38]. Suhanic [36] addressed the use of the banana envelope to show the integration of project resources and project schedule stated as early and late dates. The S-curve is based on combinations of historical projects and is popularly formed by third, fourth, or fifth degree polynomials [30]. The S curve is normalized to two basic quantities: total quantity (such as project costs) and project duration, which is widely applied to various aspects of project management [10]. Stallworthy [35] utilized S curve for "the value of work done" to a project cost control in relation to estimates of cost and time. Singh and Lokanathan [34] indicated that the application of S curves to cash flow projections can achieve an accuracy of approximately 88–97%. Barraza et al. [2] introduced a stochastic S curve as an alternative to the deterministic S curve that generated most likely budget and duration values. Blyth and Kaka [3] produced an individual S curve for a particular project using a multiple linear regression model with studies on 50 projects and 20 criteria.

Many scholars have made efforts to develop artificial intelligence (AI) models and/or establish AI systems to address various practical problems. The application of AI approaches is currently considered as an alternative approach to control project costs. In using AI

* Corresponding author.
E-mail address: tsaihsingchih@gmail.com (H.-C. Tsai).
[1] #43, Sec. 4, Keelung Rd., Taipei, Taiwan, R.O.C. 106, Dept. of Construction Engineering. Tel.: +886 2 2733 6596; fax: +886 2 2730 1074.
[2] #43, Sec. 4, Keelung Rd., Taipei, Taiwan, R.O.C. 106, Dept. of Construction Engineering. Tel.: +886 2 2730 1277; fax: +886 2 2730 1074.
[3] #43, Sec. 4, Keelung Rd., Taipei, Taiwan, R.O.C. 106, Dept. of Construction Engineering. Tel.: +886 2 2719 9999; fax: +886 2 2546 4577.

0926-5805/$ – see front matter.
doi:10.1016/j.autcon.2008.10.005

approaches, two major tasks are necessary. The first is to develop an appropriate approach and the second is to build an applicable historical data pool with relevant parameters. With AI approaches, even though the relationship between inputs and output are not identified, predicted result can still be accurately assessed. Chua et al. [9] used neural networks to assess project budget performance with 27 input factors of 22 data records. Boussabaine and Kaka [5] used a neural network approach with initial cash flow for periods to address sequential period cost flows. Lowe et al. [25] used Expert Systems to help clients manage their cash flow. Boussabaine et al. [4] used artificial neural networks to aid in the forecasting of water pipeline project cash flows. Lokmic and Smith [24] introduced back-propagation neural networks as an alternative for cash flow forecasting, in which self-organizing feature mapping was used for project clustering. Hegazy and Petzold [15] developed a dynamic project cost control model that used genetic algorithms for total cost optimization. When a project is constrained by discounted cash flows, the ant colony algorithm can be used for resource scheduling [33]. Authors of this paper contributed to developing hybrid AI tools, which were subsequently employed in AI approaches that have been successfully applied in many construction industry research topics, especially in the area of construction management [7,8].

This paper not only contributes to the application of AI learning to cash flow prediction, but also introduces an original concept of strategic project cash flow control. In the remaining sections, employed artificial intelligence approaches used to develop the project cash flow forecasting model will be introduced. The cash flow learning result will further be applied to a strategic project cash flow control to make sure the project cash flow performs within the banana envelop of the S curve. With the strategic approach, construction projects should be performed with a greater percentage chance of success.

2. Artificial intelligence approaches

2.1. K-means clustering

Many algorithms are able to identify specific domains. K-means clustering is a simple and fast approach to cluster data that starts with k centroids (seeds) that are typically generated randomly. Each data set (sample) is assigned to the cluster with the closest centroid based on Euclidean distance measurement. It is customary to set a threshold on iteration numbers to prevent excessive calculation time. Every clustering feature can be determined after repeated iteration steps. As desired cluster numbers can be set as a limitation for target convergence, a perfect convergence cannot be guaranteed. K-means usually converges in practical applications, especially in pattern recognition problems. K-means clustering is widely and commonly employed owing to its simplicity, although it does present inherent drawbacks, such as its requiring a specified setting for the optimal solution or time consumption [26].

When input data set S is composed of n points (n d-dimensional vectors), the k cluster centroids C must be satisfied with following descriptions [28]:

$$S = \{x_1\, x_2\, \ldots\, x_n\} \tag{1}$$

$$C_m \neq \Phi, \quad m = 1 \sim k \tag{2}$$

$$C_m \cap C_n = \Phi, \quad m, n = 1 \sim k, \quad m \neq n \tag{3}$$

$$\bigcup_{i=1}^{k} dataset(C_m) = S \tag{4}$$

$$d\left(x_i, C_j\right) = \left(x_i - C_j\right)^T \left(x_i - C_j\right), \quad i = 1 \sim n, \quad j = 1 \sim k. \tag{5}$$

The above definition describes each cluster as having at least one dataset. Each dataset belongs to a cluster with one-to-one relationships and every dataset must attach to a non-null cluster. The k cluster centroids are initially selected at random from S. During each iteration process, every data point x is assigned to a particular cluster set using the closest Euclidean distance measurement d. Once every data point has been assigned to a cluster, all centroids C can be re-calculated by means of all attachment points. This represents the principal concept of the K-means algorithm. K-means iterates until stable cluster centroids are found [37].

2.2. Genetic algorithm

The Genetic Algorithm (GA), which imitates parts of the natural evolution process, was first proposed by Holland [16]. GA is a stochastic search approach inspired by natural evolution that involves crossover, mutation, and evaluation of survival fitness. Genetic operators work from initial generation to offspring in order to evolve an optimal solution through generations. Each individual of a generation generates a result for the problem and is represented as a string-named chromosome. The relatively straightforward and simple implementation procedure gives GA its exceptional flexibility to hybridize with domain-dependent heuristics to create effective implementation solutions tailored to specific problems. Based on these merits, the potential of using GA in optimization techniques has been studied intensively [11]. However, simple GA is difficult to apply directly and successfully to a large range of difficult-to-solve optimization problems [29].

2.3. Fuzzy logic

Zadeh [39] first proposed Fuzzy Logic as a tool with which to describe uncertainty and imprecision. Because FL imitates the high order mode in which the human brain makes decisions in the face of uncertainty or vagueness, it provides an effective way for automated systems to describe highly complex, ill-defined, or difficult-to-analyze subjects. In general, Fuzzy Logic is composed of a fuzzifier, rule base, inference engine and defuzzifier [22]. The FL approach still has certain problems to overcome, such as membership function configuration, composition operator determination, and application-specific fuzzy rule acquisition [27]. Although FL parameters can be determined using the experience and knowledge of experts, determining these parameters in the absence of such experts remains difficult for particularly complex problems [13].

2.4. Neural networks

A Neural Network (NN) focuses primarily on computing and storing information within a structure composed of many neurons. Because NN imitate the human brain in terms of learning, recall and generalization, they are usually designed to solve non-linear or ill-structured problems [14]. The human brain is robust, fault tolerant, highly parallel and able to handle information with noise. Modeling as well as the brain does is what some AI approaches try to accomplish. A frequently used NN model is multilayer perceptron learning with error back-propagation. However, appropriate NN topology and parameters are essential for accurate problem assessment. As the optimal network topology is highly problem-oriented, an assessment with a high degree of accuracy is difficult to achieve [23]. In addition, some real world applications are hampered by lack of training techniques able to find a global optimum set of weights reliably [17].

2.5. Evolutionary fuzzy neural inference model

The Evolutionary Fuzzy Neural Inference Model (EFNIM) proposed by the authors of this paper employs Genetic Algorithms, Fuzzy Logic