Journal of Cleaner Production 109 (2015) 144–151

Contents lists available at ScienceDirect

Journal of Cleaner Production

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

Early prediction of the performance of green building projects using pre-project planning variables: data mining approaches

Hyojoo Son, Changwan Kim*

Department of Architectural Engineering, Chung-Ang University, Seoul 156-756, Republic of Korea

ARTICLE INFO

Article history:
Received 29 August 2013
Received in revised form
15 August 2014
Accepted 21 August 2014
Available online 2 September 2014

Keywords:
Green building project
Certified green building
Pre-project planning
Cost performance
Schedule performance
Data mining

ABSTRACT

Early prediction of the success of green building projects is an important and challenging issue. The aim of this study was to develop a model to predict the cost and schedule performance of green building projects based on the level of definition during the pre-project planning phase. To this end, a three-step process was proposed: pre-processing, variable selection, and prediction model construction. Data from 53 certified green buildings were used to develop the models. After balancing the data set with respect to the proportion of cases in each of the outcome categories by pre-processing, the number of input variables was reduced from 64 to 13 and 7 for cost and schedule performance prediction respectively, using the ReliefF-W variable selection method. Then, cost and schedule performance prediction models were constructed using the selected variables and four different classifiers: a support vector machine (SVM), a back-propagation neural network (BPNN), a C4.5 decision tree algorithm (C4.5), and a logistic regression (LR). The classification performance of the four models was compared to assess their applicability. The SVM models exhibited the highest accuracy, sensitivity, and specificity in predicting both the cost and schedule performance of green building projects. The results of this study empirically validated that the cost and schedule performance of green building projects is highly dependent on the quality of definition in the pre-project planning phase.

1. Introduction

Green building, a practice that is two decades old, has become more prevalent in recent years. The U.S. Environmental Protection Agency defines "green building" as the practice of creating structures and using processes that are environmentally responsible and resource-efficient throughout a building's lifecycle, from siting to design, construction, operation, maintenance, renovation, and deconstruction (U.S. Environmental Protection Agency, 2010a). Many countries have developed and adopted various green building rating systems (Zhang et al., 2013). The exemplar systems are the U.K.'s Building Research Establishment Environmental Assessment Method (BREEAM), the U.S.' Leadership in Energy and Environmental Design (LEED), Australian Green Star, Germany's Deutsche Gütesiegel Nachhaltiges Bauen (DGNB), the Japanese Comprehensive Assessment System for Building Environmental Efficiency (CASBEE), and the Korean Green Building Certification System (K-GBCS). These systems outline specific guidelines for implementing green practices into the lifecycle of a building, thus guiding the stakeholders of the building project in the green delivery of their project.

The number of buildings certified with green building rating systems has rapidly increased as a result of the rapid spread of such systems and the recognition of the benefits of green buildings, which include reduced operating costs; the creation, expansion, and development of markets for green products and services; improved occupant productivity; and optimized life-cycle economic performance (Zhang, 2014; Atlee, 2011; U.S. Environmental Protection Agency, 2010b; Shen et al., 2010). For example, the total U.S. green building market value is expected to increase from $10 billion in 2005 to between $98 and $106 billion in 2013 (McGraw-Hill Construction, 2012). This rapid growth of the green building market has raised concerns among project stakeholders about the risks involved, especially the high level of uncertainty with respect to project performance in the delivery of green building projects (Hwang and Leong, 2013; Robichaud and Anantatmula, 2011). Hwang and Leong (2013) reported that the failure rate of 39 green building projects in their survey was 33% in terms of schedule performance, more than twice that of traditional building 40 projects.

The successful delivery of green building projects is more difficult and complex than for traditional building projects, particularly

* Corresponding author. Tel.: +82 2 820 5726; fax: +82 2 812 4150.
E-mail addresses: changwan@cau.ac.kr, changwan72@gmail.com (C. Kim).

http://dx.doi.org/10.1016/j.jclepro.2014.08.071
0959-6526/

due to the additional considerations which must be taken into account in the pre-project planning phase (Zhang et al., 2011a). These include advanced simulation and analysis, higher construction standards, the contribution of multidisciplinary experts, site precautions, and knowledge of sustainable building practices (Zhang et al., 2011b; Pulaski et al., 2006). Poor or incomplete pre-project planning is at the root of failures of green building projects (Chandramohan et al., 2012; Robichaud and Anantatmula, 2011).

In previous studies, pre-project planning practice was recognized as an important contributor to the success of green building projects. Pulaski et al. (2006) reported that significant improvements in the performance of green building projects can be made by giving consideration to constructability in early design phase. Robichaud and Anantatmula (2011) highlighted the value of the involvement of a cross-disciplinary team during the pre-project planning phase in determining the success of green building projects. Swarup et al. (2011) stated that the involvement of the contractor in the early design phase is vital to achieve success in green building projects. Zhang et al. (2011a) acknowledged the importance of engaging the stakeholders and encouraging them to have efficient communication during the planning and design phases to their successful delivery of green building projects. Although previous studies have emphasized the importance of specific pre-project planning practices, none have explored an empirical relationship between the level of definition in the pre-project planning phase and the performance of green building projects. Furthermore, the predictability of success by taking the implementation of pre-project planning activities into account has not been validated empirically. Early prediction of the performance of a green building project is critical for project stakeholders to make decisions that could influence project success.

The aim of this study is to develop a model to predict the cost and schedule performance of green building projects based on the level of definition during the pre-project planning phase. The rest of this paper is organized as follows: Related studies on the influence of pre-project planning on the performance of building project are discussed in Section 2. Data collection and the characteristics of the data set are outlined in Section 3. The research methodology is presented in Section 4 and the experimental results are given in Section 5. Finally, conclusions are presented in Section 6, which also highlights the contribution from this research and mentions areas that merit future research.

2. Related studies

Considerable research has been devoted to investigating the influence of pre-project planning on the performance of building project. Some researchers have applied statistical methods, such as mean value difference comparison and statistical regression analysis, to validate the relationship between pre-project planning and the project performance (for example, Wang and Gibson, 2010; Wang, 2002; Cho and Gibson, 2001).

More recent studies have applied data mining techniques to explore the relationship between pre-project planning phase and the financial performance of building project. Wang et al. (2012) used a support vector machine model to examine the relationship between sub-scores from three different sections of the Project Definition Rating Index (PDRI) and the cost and schedule performance of 92 building projects. The models had a predictive accuracy of 92% and 81% for cost and schedule performance respectively. Son et al. (2012) developed a hybrid predictive model to understand the impact of pre-project planning on cost performance, combining principal component analysis (PCA) with a support vector regression (SVR) model. Values for 64 PDRI variables and cost performance data from 84 commercial building projects were used to develop the model. The cost performance values predicted by the PCA–SVR method were very close to the actual values (mean absolute percentage error less than 10%). All of the previous studies which have shown that the level of project definition influences the performance of building projects focused on traditional construction projects. None have investigated the predictability of the success of green building projects based on the level of definition in the pre-project planning phase. The question then arises whether the level of project definition during the pre-project planning phase influences the performance of green building projects.

3. Data collection

3.1. Interview design

Data collection was based on a face-to-face interview followed by a questionnaire for 53 respondents. The survey respondents included the key project stakeholders (i.e. architect, developer, or project manager) for each project. These persons had access to detailed information about the progress of the project during the pre-project planning phase and had responsibility for the project at that time. The questionnaire was designed to measure the level of definition in the pre-project planning phase and the cost and schedule performance of each project.

In this study, the 64 scope definition elements in the PDRI for buildings were employed as the independent variables to measure the level of definition in the pre-project planning phase of each project. The PDRI for buildings, developed by the Construction Industry Institute (CII, 1999) provides a 64-item checklist that encompasses all the project activities during the pre-project planning phase. The elements in the checklist were developed with the participation of more than 100 industry practitioners and were developed with consideration of their impact on overall project performance (Cho and Gibson, 2001). These practitioners included engineers, architects, and other industry professionals directly involved in planning and executing building projects. Hence, the checklist comprises comprehensive and extensive elements that are most suitable for measuring the level of definition in the pre-project planning phase of each project (Cho and Gibson, 2001).

In the PDRI for building projects, the 64 variables are divided into 11 categories, which in turn are grouped into three sections: basis of project decision (18 variables), basis of design (32 variables), and execution approach (14 variables). The variables and their groupings are listed in Appendix I. Detailed descriptions of the variables are published by the CII (1999). To collect information on the level of definition in the pre-project planning phase for each project, the interviewer asked the respondents to evaluate each of the 64 variables on a five-point Likert-type scale. The scale ranged from 1 (complete definition) to 5 (incomplete or poor definition).

Respondents were also asked to provide data about the actual recorded cost and duration at the time of project completion, as well as the budgeted cost and planned duration estimated at the time that the decision was made to proceed with the project. From this data, a project was classified as a cost failure if its actual cost exceeded budgeted cost, or as a schedule failure if its actual duration exceeded planned duration. Otherwise, the project was considered successful.

3.2. Data profile

The survey targets were 53 certified green buildings constructed in South Korea over the past five years from 2008 to 2012. These