A hybrid approach using data envelopment analysis and case-based reasoning for housing refurbishment contractors selection and performance improvement

Yi-Kai Juan *

Department of Architecture, National Taiwan University of Science and Technology (NTUST), 43, Section 4, Keelung Road, Taipei 106, Taiwan

ARTICLE INFO

Keywords:
Data envelopment analysis (DEA)
Case-based reasoning (CBR)
Housing refurbishment
Contractors selection
Performance improvement

ABSTRACT

The refurbishment market has grown greatly in the last decade. Relevant projects are becoming increasingly more demanding in the construction industry due to the emphasis on sustainability. Most refurbishment works, however, involve a higher level of risk and uncertainty, as well as more complex coordination than new buildings. These characteristics are likely to cause asymmetric information between contractors and tenants in a refurbishment process and thus affect customers’ satisfaction and project performance. This study proposes a systematic decision support approach to solve refurbishment asymmetric information problems by using case-based reasoning (CBR) and data envelopment analysis (DEA). The PZB model of the service quality and fuzzy sets are applied to support the DEA operation. With this intelligent approach, tenants can select an optimal refurbishment contractor according to their customization needs and contractors can find out their inefficient factors to improve their business competitiveness. The proposed hybrid decision support approach is expected to be useful for both tenants and contractors in those developed countries or regions with high refurbishment needs when they face similar problems.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

With the change in social structures, an increasing awareness of the need for sustainability, as well as the high expectation of living environments, the refurbishment market has faced blooming needs worldwide in the last decade (Egbu, 1997). Global organizations have invested plentiful resource in creating sustainable refurbishment environments. In some developed countries, the output value of refurbishment almost reaches a half of the total construction output (CCCIS, 1996). The refurbishment issue undoubtedly will be a major issue for sustainable development in the construction industry.

Refurbishment projects are usually characterized by complex, small-scale and highly labor-intensive renovation tasks (Daoud, 1997; Okoroh & Torrance, 1999). It is generally accepted to be more difficult to manage and needs greater coordination than new building, due to a higher level of risk and uncertainty (Quah, 1992). Some research reveals that one of the severest challenges of refurbishment projects is asymmetric information between contractors and tenants in a refurbishment process (Holm, 2000). Inadequate refurbishment knowledge of tenants, ineffective communication, and unstable service quality of contractors usually affect customers’ satisfaction and project performance.

To solve the above-mentioned problems, what tenants need is a healthy contractor who can provide the most advantageous refurbishment tender, offer reliable service, and possess integrated capability. For contractors, they also need an objective system to discover inefficient factors leading to unstable service quality to improve their business performance. Traditional construction management has explored several studies regarding how to evaluate qualified contractors in a bidding process (Alsugair, 1999; Chua & Li, 2000), as well as how to improve construction project performance (Cottrell, 2006; Kagioglou, Cooper, & Aoud, 2001). In refurbishment, however, there has been far less research in this domain because most housing refurbishment project evaluations are based on intangible services, word-of-mouth referrals (Holm, 2000), or are based on customers’ intuitive judgments (Okoroh & Torrance, 1999). In addition, most refurbishment works are conducted by unskilled “cow boy” operators, which has increased numerous management difficulties (Ranaweera & Prbhu, 2003). Therefore, an effective and structured contractor selection process for tenants and the self-examination of refurbishment performance by contractors need to be simultaneously developed.

A hybrid approach combining case-based reasoning (CBR) and data envelopment analysis (DEA) is built to explore the two-way asymmetric information between tenants and contractors. By using the CBR technique, tenants can retrieve the past refurbishment cases of candidate contractors from a database according to individual housing conditions and customized refurbishment needs. Contractors need to be simultaneously developed.
needs. The DEA technique, supported by the PZB model of service quality and fuzzy sets, is then used as a case revision process of CBR to evaluate each candidate contractor’s efficiency to determine an optimal contractor. A sensitivity analysis (SA) technique is finally applied to disclose inefficient factors to improve contractors’ business performances. 254 housing refurbishment projects, conducted by 62 companies, are collected for a pilot test. It is expected that this approach can make a contractor selection process much more desirable and contractor’s performance improvement strategies more convincing, which will significantly increase customer’s satisfaction and contractor’s competitiveness.

2. Service quality in the refurbishment industry

Refurbishment works are somewhat different from new construction projects. Special characteristics of housing refurbishment contain the site-driven works undertaken in an existing building (Daoud, 1997), the intensified uncertainty (Clancy, 1995), the long turn-round time (Aikivuori, 1996), and many simultaneously operating workers in a restricted space. These unique characteristics render it more difficult to standardize the delivered service, compared to new construction, and the outcome of the refurbishment performance will be more dependent on contractor’s capability and experience. There is likely to be an increasing level of innovative approaches to procuring refurbishment projects (Egbu, 1997).

Refurbishment can be categorized as a service industry (Holm, 2000). Providing what customers expect, such as customized products or services, is key to reach customer satisfaction in service management. In a refurbishment process, how to select healthy contractors will be tenants’ most important concerns. In the past, however, contractors only had high interest in speculatively making great profits to keep continuous competitive advantage in the refurbishment market. There is a gap between these two parties and it has easily become a source of asymmetric information. Parasuraman, Zeithaml, and Berry (1985) have proposed a service quality model indicating customers’ quality perceptions are influenced by four distinct gaps, including the perception gap, the specification gap, the delivery gap, and the communication gap. Five dimensions, including “Tangibles”, “Reliability”, “Responsiveness”, “Assurance”, and “Empathy”, to evaluate service quality, have been developed (Zeithaml, Berry, & Parasuraman, 1988). The innovative concept of service quality is useful to balance the forces between the “pull” of tenants and the “push” of contractors. Many contractors have been conscious of a significant relation between good customer service and the creation of competitive advantages (Rapert & Wern, 1998), because most refurbishment projects, like other service works, come from word-of-mouth referrals. Contractors who want to take advantage of asymmetric information to acquire profits do not necessarily ensure their long-term advantages (Torbica & Stroh, 2001). Therefore, the service quality is crucial for improving asymmetric information problems between tenants and contractors.

3. CBR and DEA

3.1. Case-based reasoning (CBR)

CBR is a problem solving technique re-using past cases and experiences to find a solution to problems (Juan, Shih, & Perng, 2006). The central tasks involved in CBR methods are to identify the current problem situation, find a past case similar to the new one, use that case to suggest a solution to the current problem, evaluate the proposed solution, and update the system by learning from this experience (Shin & Han, 2001). This technique is an alter-native to an expert system, which is based on rule-based reasoning. Essential tasks to make a typical CBR system are: (1) retrieve one or a small set of the most similar cases according to the problem specification; (2) adapt to the new situation by revising former solutions; (3) retain the new case and solution as part of history for future retrievals (Jeng & Liang, 1995).

3.2. Data envelopment analysis (DEA)

In the scientific field, many studies of the use of DEA for the evaluation of effectiveness of various objects are presented (Rimkuviene, 2004). DEA is an approach for comparing the relative efficiency of decision making units (DMUs). Usually the investigated DMUs are characterized by a vector of multiple inputs and multiple outputs. To aggregate information about input and output quantities, DEA makes use of fractional and corresponding linear programs to measure the relative performance of DMUs (Wang, 2005). DEA is based on the economic notion of Pareto optimality. A given DMU is not efficient if some other DMU can produce the same amounts of output with less of some resources and not more of any other (Al-Shammari, 1999).

3.3. A hybrid DEA-based CBR

The case revision process is usually implicit for the traditional application of CBR because it may involve the user’s subjective judgment to adapt retrieval cases. Some research even regards this as a ‘black box’ (Chua, Li, & Chan, 2001). This defect might limit the accuracy of the CBR operation, as well as decrease the robustness of applications. In this study, a two-stage decision support model is developed by adopting a hybrid DEA-based CBR to select potential refurbishment contractors and then evaluate their performances. With this hybrid approach, DEA is integrated into the revision process of CBR, which can enhance the robustness and persuasiveness of the CBR operation.

4. A systematic decision support approach

This study attempts to integrate the DEA into the CBR process to evaluate the contractor’s performance. In this model, as shown in Fig. 1, CBR is used to retrieve past refurbishment cases which are most consistent with customers’ current housing conditions and refurbishment requirements. Any case can be represented in a variety of forms and categorized into relevant attributes so that it can be easily extracted from the case base during the retrieval stage.

4.1. Case representation

In the case representation process, seven attributes judged with regard to the search goal are defined: “Housing Type” (condominium, town-house, cottage, low-rise apartment, high-rise apartment, etc.), “Refurbishment type” (indoor redecoration, facility maintenance, building extension, material replacement, etc.), “Building year”, “Location”, “Area”, “Cost”, and “Schedule”. In the retrieval process, the most common case retrieval techniques, the indexing tree concept and the nearest neighbor method, are introduced.

4.2. Case retrieval: the combination system of indexing and nearest neighbor

Case indexing involves assigning indices to cases to facilitate their retrieval (Tseng, Chang, & Chang, 2005). This approach generally generates a tree-type structure to organize the cases in mem-
The nearest neighbor approach, on the other hand, is a simple and straightforward method to assess the similarity between stored cases and the newly input case, as shown in Eq. (1), based on matching a weighted sum of attributes (Chang, Cheng, & Su, 2004). That is, every attribute in the input case is matched to its corresponding attribute in the stored case, and the degree of match of each pair is computed using the matching function (Choy, Lee, Lau, & Choy, 2005).

\[
\text{Similarity}(T, S) = \sum_{i=1}^{F} \left( W_i \times \left( 1 - \sqrt{\frac{T_i - S_i}{T_i}} \right)^2 \right)
\]  

(1)

where \( W_i \) is the weight of feature \( i \), \( T \) is the target case, \( S \) is the source case, \( F \) is the number of attributes in each case, and \( i \) is an individual feature from 1 to \( F \).

To overcome a traditional problem in which the nearest neighbor approach has difficulty in deciding a set of attribute weights to accurately retrieve cases (Park & Han, 2002), this study uses the analytic hierarchy process (AHP) method to aid tenants in determining each attribute weight according to their preferences.

Some studies reveal that combining these two retrieval methods is effective and advantageous during the retrieval process, especially for dealing with complex cases or a large number of

Fig. 1. Evaluation model combining CBR and DEA.
cases (Shin & Han, 2001). As shown in Fig. 2, a combination of inductive indexing and nearest neighbor methods supplementing each other is a promising solution for finding a satisfactory match case.

**Fig. 2.** Case retrieval process.
4.3. Case revision: DEA supported by PZB model and fuzzy sets

In the case revision process, DEA is adopted to evaluate the performance of potential contractors retrieved from the case base. The Charnes, Cooper and Rhodes (CCR) ratio model is applied in conducting DEA (Charnes, Cooper, & Rhodes, 1978). Data on output measures are based on five dimensions of the service quality theory, named the PZB model, which is useful to evaluate customers’ quality expectation. Input measures, involving in affecting the performance of output, are obtained from interviewing 12 experienced experts and contractors. An in-depth interview focused on identifying the factors in a refurbishment process that may affect the project satisfaction. Therefore, input measures contain X1 (Schedule change), X2 (Budget change), X3 (Quality defect), X4 (Contractor experiences), and X5 (Conflicts and complaints), while output measures contain Y1 (Tangibles satisfaction), Y2 (Empathy satisfaction), Y3 (Reliability satisfaction), Y4 (Responsiveness satisfaction), and Y5 (Assurance satisfaction).

To conduct DEA appropriately, these inputs and outputs need to be transformed into a quantitative data representation. For output, a score from 1 to 100 is defined to describe the satisfaction degree of each measure. Scores are determined by a set of questionnaires answered by concerned tenants. The higher the score, the more satisfied the tenant is with the project. For each input measure, five impact levels are defined to describe the influence of each measure’s performance. Table 1 and Fig. 3 are samples of a description for the first input measure, “Schedule change”. To quantify impact levels of these measures, and to reflect the actual considerations from different experts, a fuzzy sets theory is used to describe the inherent imprecision and ambiguity associated with the description of impact levels for evaluating input measures (Seo, Aramaki, Hwang, & Hanaki, 2004). Since different experts have significant discrepancies in judgment on impact levels of each input, a fuzzy number (\(\tilde{E}_{ij}^{k}\)) is built to help characterize the uncertainty, which is inherent in a given set with a degree of membership. Each fuzzy set is defined by its membership function, \(\mu_{i}(x)\), associated with a numerical value in the interval [0, 1] (Zhang, Tam, & Shi, 2003).

Table 1
Impact level description of “Schedule change”

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
<th>Fuzzy number</th>
<th>Defuzzification</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>The schedule change ratio is less than 3%</td>
<td>(1.00, 1.00, 3.00)</td>
<td>1.67</td>
</tr>
<tr>
<td>4</td>
<td>The schedule change ratio is about 5%</td>
<td>(2.40, 3.70, 5.00)</td>
<td>3.70</td>
</tr>
<tr>
<td>3</td>
<td>The schedule change ratio is about 10%</td>
<td>(4.20, 5.50, 7.00)</td>
<td>5.57</td>
</tr>
<tr>
<td>2</td>
<td>The schedule change ratio is about 15%</td>
<td>(6.00, 7.40, 8.60)</td>
<td>7.33</td>
</tr>
<tr>
<td>1</td>
<td>The schedule change ratio is more than 20%</td>
<td>(8.00, 10.00, 10.00)</td>
<td>9.33</td>
</tr>
</tbody>
</table>

The fuzzy numbers are distributed from 1 to 10 score. The higher the score, the worse the performance.

Fig. 3. Fuzzy number of ‘Schedule change’.

Fig. 4. Fuzzy numbers of impact levels for each input measure.
denote lower, medium, and upper values of the impact level, respectively.

The fuzzy set that describes a linguistic value sometimes has to be expressed by a crisp value to illustrate the impact levels of quantitative criteria, which means defuzzification. The defuzzification method, as shown in Eq. (3), adopted here is the center of gravity method (CGM) (Bojadziev & Bojadziev, 1997). If the fuzzy set is a triangular membership function, it can be simplified as

<table>
<thead>
<tr>
<th>Table 2</th>
<th>DEA for evaluating contractors' efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMU</td>
<td>Score (%)</td>
</tr>
<tr>
<td>C05</td>
<td>96.51</td>
</tr>
<tr>
<td>C21</td>
<td>83.15</td>
</tr>
<tr>
<td>C18</td>
<td>79.70</td>
</tr>
<tr>
<td>C07</td>
<td>100.00</td>
</tr>
<tr>
<td>C55</td>
<td>67.25</td>
</tr>
<tr>
<td>C59</td>
<td>49.37</td>
</tr>
<tr>
<td>C41</td>
<td>63.13</td>
</tr>
<tr>
<td>C12</td>
<td>100.00</td>
</tr>
<tr>
<td>C59</td>
<td>83.04</td>
</tr>
<tr>
<td>C41</td>
<td>47.24</td>
</tr>
<tr>
<td>C33</td>
<td>92.02</td>
</tr>
<tr>
<td>C08</td>
<td>52.99</td>
</tr>
<tr>
<td>C47</td>
<td>34.52</td>
</tr>
<tr>
<td>C36</td>
<td>77.49</td>
</tr>
<tr>
<td>C31</td>
<td>84.35</td>
</tr>
<tr>
<td>C11</td>
<td>55.41</td>
</tr>
<tr>
<td>C39</td>
<td>64.47</td>
</tr>
<tr>
<td>C50</td>
<td>65.09</td>
</tr>
<tr>
<td>C60</td>
<td>35.87</td>
</tr>
<tr>
<td>C44</td>
<td>94.47</td>
</tr>
<tr>
<td>C31</td>
<td>83.01</td>
</tr>
<tr>
<td>C23</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Fig. 5. Sample of CBR initial results.
shown in Eq. (4). Fig. 4 illustrates defuzzification of fuzzy sets of each input measure. Each contractor can therefore be evaluated by score according to the criteria of Fig. 4.

\[
NF = \frac{\int x \cdot \mu(x) \, dx}{\int \mu(x) \, dx}
\]  

(3)

\[
NF = \frac{(LE_{ij}^k + ME_{ij}^k + UE_{ij}^k)}{3}
\]  

(4)

5. Case study

5.1. Data collection

Two hundred and fifty four refurbishment projects in five major cities of Taiwan were collected and refined into a database. The satisfaction questionnaire was carried out by each concerned tenant who finished a refurbishment project in the study year. For DEA, data collection of output came from these satisfaction questionnaires. Input data collection was provided by 62 refurbishment contractors responsible for these projects. However, not all initial input and output can be used for operating DEA because of a limited number of DMUs. Therefore, a statistical test was carried out to examine the correlation among the input and output. After this procedure, the acceptability of input and output of DEA is confirmed.

5.2. Pilot test

An actual house-owner and his refurbishment requirements are proposed to test the DEA-based CBR. The dwelling, with total indoor area 130 M², is a unit of a high-rise apartment constructed in 1996 in Taipei city, a northern area of Taiwan. The tenant has a limited budget, US$48,500, and a constricted construction schedule of 70 days for conducting a whole indoor redecoration project.

By using the proposed indexing and the nearest neighbor method of CBR, the result of suggested initial solutions can be retrieved from a database. 40 projects, conducted by 22 contractors, with a similarity value of at least 75% are retrieved as a sample shown in Fig. 5. Fig. 5 also expresses fuzzy value scores of each project’s input and output data.

CBR retrieval initial results need to be further revised by adopting DEA. 22 contractors are regarded as DMUs being evaluated for efficiency. The result, in Table 2, apparently reveals that contractors “C07”, “C12”, and “C23” are relatively efficient in refurbishment performance, while the other 19 contractors (86.4%) are
relatively inefficient. A matrix analysis concerning the efficiency (X-axis) and project satisfaction (Y-axis) of these contractors is depicted in Fig. 6. In the matrix, the contractor “C23” seems to be the optimal choice for the tenant, with higher satisfaction and the same efficiency. Contractors who locate in the II and III quadrants should find effective strategies to move forward to the I quadrant to increase their customers’ satisfaction and business competitiveness. No contractors are located in IV quadrant and it is reasonable for DEA because an efficient DMU should not have lower output.

Fig. 6 also provides insights for inefficient contractors. In terms of strategic planning, they can be aware of their competitive situations, compared to their rivals, in a severely competitive refurbishment market. A sensitive analysis (SA) technique is further used to help contractors to find out which dimensions of input and output should be improved urgently.

Take “C47” and “C60” who are the two least efficient contractors by test. The result, as shown in Fig. 7, shows that the “Quality defect” for input and “Empathy satisfaction” for output are the most important considerations for the “C47” contractor to make improvement decisions. In this case, the contractor should strive to establish a healthy quality management system or arrange for a responsible site supervisor to control the construction work to reduce unnecessary defects. In addition, he has to pay more attention to his customers with empathy concerns to increase satisfaction, such as more convenient service and thoughtful, proactive tenant problem solutions. Similarly, the “Schedule change” of input and “Tangibles satisfaction” of output have high priorities for the “C60” contractor if he expects to polish his business performance. In the near future, he should plan his project schedules more cautiously to avoid work delays and provide more tangible services, such as appropriate equipment, neatness in appearance, etc., to improve his business efficiency. Both “C47” and “C60” can acquire 80–100% efficiency improvement (add up input and output) if they are willing to give weight to the most urgent and sensitive dimensions.

The decision support feedback is truly essential for contractors to diagnose their business defects, know their competitive position in the refurbishment market, and, most important of all, make improvement strategies to promote their business performance. It is meaningful, from a strategic decision making viewpoint, to put emphasis on those urgent and influential factors to create competitive advantages especially when under limited resource constraints.

6. Conclusions and suggestions

With limited land usage and being aware of sustainability, refurbishment undoubtedly will become an important market in the construction industry of most developed countries. Traditional refurbishment processes indeed cause asymmetric information between contractors and tenants, as well as, induce large numbers of needless complaints and disputes, which may affect customer’s satisfaction and project performance.

Currently most refurbishment contractor selections are usually based on word-of-mouth referrals or tenant’s intuitive judgments. It may lead to a vicious cycle of asymmetric information if, unfortunately, a tenant is not able to select an appropriate contractor. Some tenants even hold attitudes such that they have no choice but to conduct refurbishment work at a heavy price. For contractors, many refurbishment jobs are undertaken by open shop workers or by a homeowner without systematic trainings. The supply of unskilled “cow boy” operators and unstable service quality have damaged companies’ market positions. Contractors need a self-examination mechanism to improve their competitive advantages.

This study has presented a promising evaluation approach for solving two-way asymmetric information problems. With this system, a tenant can select an optimal contractor and a contractor can be provided with feedback information to improve business performance. The proposed approach adopts the DEA-based CBR technique, and takes the PZB model of service quality and fuzzy sets into account to enhance the DEA operation. A comprehensive integration of these concepts has made the approach more robust and effective. The CBR limitations and DEA deficiencies are also addressed. It is expected that the approach can lead the refurbishment industry forward to sustainable development and would be useful for those developed countries or regions, with a great quantity of refurbishment needs, when they face similar problems.

References


Please cite this article in press as: Juan, Y. K., A hybrid approach using data envelopment analysis and case-based reasoning ..., Expert Systems with Applications (2008) xxx–xxx.


