Approximate Life Cycle Assessment Using Case-Based Reasoning for the Eco Design of Products

Myeon-Gyu Jeong*, James R. Morrison, Hyo-Won Suh

Abstract—Most manufacturing enterprise should perform environmental impact assessment throughout the entire life cycle of their products due to the increased environmental consciousness of customers and the introduction of strict environmental regulations. Life cycle assessment (LCA) is a systematic method of analyzing the environmental load of a product and evaluating potential environmental effects. It can be useful for eco design and for dealing with environmental regulations. However, the LCA process generally requires considerable time and money to collect relevant life cycle inventory (LCI) data and information. Usually, enterprises develop a new product by revising or reusing a similar previous product. Repeating the LCI data collection whenever they develop a new product is very cumbersome and will increase lead time. However, if we estimate the eco impact values based on the previous design, then the efforts related to LCI data collection and its attending complex computational procedures are reduced. Although various efforts have been made on the streamlined LCA in an effort to overcome these limitations, the result is still unsuitable for practical eco design. We therefore propose an approximate LCA method using case-based reasoning (CBR) for a rapid and convenient environmental evaluation in product development. For the approximate LCA, we developed function behavior structure environmental effect (FBSE) representations and a creative similarity measurement based on FBSE for a clear and consistent CBR process. A geometry attribute based linear modeling algorithm of eco impact is proposed to replace complicated LCA procedures, and genetic algorithms are used to search for optimal solutions to satisfy the proposed model. A case study involving an upstream process of a vehicle air purifier fan confirms that the proposed CBR for LCA method can be effectively applied to eco product design.

I. INTRODUCTION

In recent years, an increase in environmental consciousness and the introduction of strict environmental regulations have challenged designers to consider the environmental performance of products together with traditional design objectives in the early stages of design. Many companies are therefore striving to obtain green certification as they endeavor to cope with environmental legislations, particularly with regard to WEEE (waste of electrical and electronic equipment), RoHS (restriction of hazardous substances directive) and EuP (energy using product) [1].

Life cycle assessment (LCA) is now the most representative and sophisticated tool for analyzing and quantifying the consumption of resources and their environmental impact throughout the entire life cycle of a product. It has been standardized by the ISO 14040–3 series for practical use in the enterprise [2]. Fig. 1 describes general LCA stages and applications.

![Figure 1. General LCA stages and applications](image_url)

In general, most enterprises develop new products by revising or reusing the similar previous product. Environmental impact assessment is required prior to detailed design to establish the eco product development strategy. However, repeating the LCA analysis for a new product that is similar to others already studied is very cumbersome and increases lead time. Although various studies with advanced LCA techniques have been conducted in an effort to overcome these drawbacks, the result is still unsuitable for practical design work.

The advanced LCA methods can be classified into three types. The first type is matrix-based qualitative or quantitative assessment [2], [3]. This type, though useful in the early design stage because of its simplicity, is more likely to reflect the subjective views of designers. The second type involves partial elimination of complicated LCA procedures or product attributes, though it has no systematic means of selecting the concerned life stage and attributes [4]. The third type involves the substitution of an artificial neural network or multi-regression equation for complicated LCA processes [5], [6]. One of drawback of these algorithms is that they are limited only to a specific impact category, such as climate change (CO₂), radiation etc. Also, their predicted eco value is not generally accurate enough to be used for practical design.

However, if we estimate LCI data of a new product from similar previous products, then the effort related to LCI data...
collection and complicated computational procedures may be
minimized. If the estimate is sufficiently accurate, the eco
product development strategy can be established based on
these results.

We therefore propose a procedure on CBR for LCA that
overcomes the limitations of previous studies. The CBR is
used as an indirect quantitative assessment method in early
design stages. The FBSE expression is suggested to develop
an effective case memory structure. A similarity measurement
based on FBSE is developed for various value types, and
k-medoids clustering algorithm is adopted to support efficient
case retrieval. A geometry attribute based linear modeling
algorithm for case adaptation is proposed to replace
complicated LCA procedures, and genetic algorithms are used
to search for optimal solutions that satisfy the proposed model.

The rest of the paper is organized as follows: In Section II,
we give an overview of the proposed method, a detailed
description of CBR for LCA and relevant issues. In Section III,
we provide a case study of the proposed method. Finally, in
Section IV, we make concluding remarks and discuss future
research issues.

II. CBR FOR LCA

A. Overview of the Proposed Method

Case-based reasoning is an approach to problem solving
that makes use of a database of previously solved problems
when solving a new problem [7]. In general, most design
gineers fail to understand the LCA process completely.
Even the designer is familiar with the LCA process, it may
require a great deal of time. CBR allows the designer to
propose solutions to problems quickly, even if the designer is
not accustomed to the field of environmental assessment.
Therefore, CBR is adopted as the framework to quantitatively
approximate LCA values in this study.

The CBR system is comprised of case representation, case
memory organization, case recall, and case adaptation [8]. The
CBR process for LCA in this study includes a creative
algorithm for similarity measurement, case memory
organization and case adaptation while following the typical
process of CBR. Fig. 2 describes the overall process of the
proposed method.

![Figure 2. The framework of CBR for LCA](image)

The first step is a case memory organization containing
both previous cases specifications and their eco impact results.
All components for each product in the database should be
functionally decomposed, and represented according to their
behavior, structure and eco impact value along with their
specifications. These previous cases should be clustered
according to their environmental characteristics for efficient
case retrieval.

The new problem P (our new product), requires a
functional decomposition and FBSE representation in the same
way, but it has no eco impact values. Its eco values will be
estimated. Each function, behavior and structure attribute will
be used as indices for case retrieval. In case retrieval, the
nearest cluster to P should be chosen first by an appropriate
similarity measurement. Once an appropriate cluster is
selected, among the cases in that cluster, some cases near P
will be used as adaptation cases.

We next conduct case adaptation, assuming that the
adaptation cases and P have, in some sense, similar
environmental characteristics. Thus, if we derive the
relationship between geometric attributes and eco impact values
in the adaptation cases, we can use these to estimate the
eco impact value of P. For this, the eco impact value is
modeled as the sum of the product of the adaptation attributes
and an optimal k-set. This optimal k-set will be obtained by
various mathematical methods such as multi regression and
genetic algorithm. It can be used to estimate not only the eco
impact of the adaptation cases but also P. Finally, the
estimated eco impact of P can be saved as a new case and it
can be used as a known case for another new problem. In this
way, our method overcomes the drawbacks of the previous
streamlined LCA. More details of above issues will be
discussed further in the following sections.

B. An FBSE expression for CBR system in LCA

The FBSE model provides the basis for classifying the
information within a case as a function, behavior and
structure [7]. Each case is represented in terms of a
multilayered representation that expresses the function,
behavior, and structure of the design entity.

We propose our own FBSE expression which enables a
general FBSE expression method to be used in LCA. Every
individual case has function, behavior, structure and
environmental effect, and each element is defined as follows:

- **Function**: The purpose of the design
  (e.g. the purpose of a fan is to move the air)
- **Behavior**: Principle to achieve intended function
  (e.g. propeller fan is a kind of fan to move the air)
- **Structure**: Description of the physical characteristics of the object
  (e.g. geometry size, material, color)
- **Environmental Effect**: Description of the environmental effect of a product
  (e.g. climate change, ozone layer)

Each case is described by a set of features and each feature
takes on a value. The features define the vocabulary for
describing the previous designs, and the values identify the
information specific to one design case [7].

A function definition developed by Hirtz et al., [9] is
chosen as the function terminology. Behavior and structure
terminology is described with standard or general terms.
Environmental effect is classified into 11 impact categories
(carcinogens, resp. organics, resp. inorganics, climate change,
radiation, ozone layer, acidification/eutrophication, land use,
ecotoxicty, minerals and fossil fuels) by eco-indicator 99
methods as in [10]. With this FBSE expression, similarity
measurement will be clear and consistent. Also the case
reasoning process and the accuracy of the CBR process will be improved.

C. Case indexing

If we classify all cases in the case memory into several clusters with similar environmental characteristics, then case retrieval will be faster. To do this, we cluster every case using the vantage based case indexing mechanism of Tsai, et al [11]. To cluster with the vantage base case indexing mechanism, we first calculate the similarity between all cases and compose an n-cluster with a representative case (called a medoid) as their center. After that, by comparing the distance between each medoid and P, we can find the most similar cluster and cases. Fig. 3 shows a diagram of the vantage based case indexing mechanism.

![Figure 3. The vantage based case clustering](image)

In Fig. 3, r1 and r2 are the maximum distance of cluster 1 (c1) and cluster 2 (c2) respectively. c2 and c3 are the medoids of c1, (v1) and c2, (v2) respectively.

In our study, we used the k-medoids clustering algorithm as our clustering method because our cases have both numeric and nonnumeric attributes. The k-medoids algorithm is a clustering algorithm related to the k-means algorithm and the medoid-shift algorithm. In contrast to the k-means algorithm, k-medoids chooses datapoints (what we call cases) as centers and work with an arbitrary matrix of distances between datapoints [12][13]. The distance calculation method for clustering is described in the next section.

D. Similarity measurement

Environmental characteristics mostly depend on the structure attribute. Structure attributes largely depend on environmental characteristics all together. However, these attributes have numeric or nonnumeric values depending on their type. Therefore, the following similarity measurements are developed to address both kinds of attribute.

1) Similarity of functional attribute

For functional attributes, we can use the function definition of Hirtz et al [9]. The functional basis of Hirtz provides a hierarchical structure of representative functions and their correspondents. Fig. 4 describes a part of their functional hierarchy. This functional definition allows a nonnumeric type. Therefore, we can obtain the functional similarity between two cases in a similar manner as tracing the degree of kinship as follow (1):

\[ f_{\text{sim}}(c_i, P) = 1 - f_{\text{dist}}(c_i, P) = 1 - \frac{d_{i1}^2 + d_{i2}^2}{2 \times \text{max} \cdot \text{dist}^2} \]  

(1)

(where, \( f_{\text{sim}}(c_i, P) \) is degree of similarity between case1 and P; it is a real number between 0 and 1)

2) Similarity of behavioral and structural attribute

Most behavioral attributes can be expressed as nonnumeric values, and structural attributes may have both nonnumeric and numeric value forms. We provide appropriate similarity measurement for both cases.

2.1) Non-numeric value

For nonnumeric values such as color or fan type, the cosine similarity of the description words can be used to estimate similarity as follows:

\[ \text{sim} = \cos(\theta) = \frac{A \cdot B}{|A||B|} = \frac{\sum (A \times B)}{\sqrt{\sum (A)^2} \times \sqrt{\sum (B)^2}} \]  

(2)

where, \( A_i \) and \( B_i \) are sets consisting of the description words, or term, for case A and B, respectively. For example, if the material of case A is “galvanized sheet steel”, case B is “aluminum rolled” and the new problem P is “cold rolled sheet steel”, then the term for each material can be express as follows:

\[ A = \{\text{galvanized, sheet, steel}\} \]

\[ B = \{\text{aluminum, rolled}\} \]

\[ P = \{\text{cold, rolled, sheet, steel}\} \]

The number of unique terms for A, B and C are then, \( T_A = 3, T_B = 2, T_P = 4 \), respectively. Since

\[ A \cup P = \{\text{galvanized, cold, rolled, sheet, steel}\} \]

\[ B \cup P = \{\text{aluminum, cold, rolled, sheet, steel}\} \]

we have that the number of terms in these unions (not necessarily unique) are \( N_{AUP} = 5, N_{BUP} = 5 \). Letting \( CT_{AUP} \) be the number of common terms between A and P,

\[ CT_{AUP} = (T_A + T_P) - N_{AUP} = (3 + 4 - 5) = 2 \]

\[ CT_{BUP} = (T_B + T_P) - N_{BUP} = (2 + 4 - 5) = 1 \]

The cosine similarity between case A and P, \( \text{sim}_{AP} \), is thus:

\[ \text{sim}_{AP} = \frac{CT_{AUP}}{\sqrt{T_A \times T_P}} = \frac{2}{\sqrt{3 \times 4}} = 0.58 \]

\[ \text{sim}_{BP} = \frac{CT_{BUP}}{\sqrt{T_B \times T_P}} = \frac{1}{\sqrt{2 \times 4}} = 0.35 \]

Therefore, case A is more similar to P than is case B.

2.2) Numeric value

i) Point matching function

A point matching function can be chosen for the case of numeric value with a maximum and minimum value. In this case, the similarity follows the distribution as Fig. 5; it is given as

\[ \text{sim} = \sqrt{1 - \frac{d_{\text{min}}^2}{d_{\text{max}}^2}} = \sqrt{1 - \left(\frac{|f_i - f_j|}{f_{\text{max}} - f_{\text{min}}}\right)^2} \]  

(3)
where, $f_i^P$ is $i$-th attribute value of $P$ and $f_i^A$ is $i$-th attribute value of case $A$, and $f_{\text{max}}$ and $f_{\text{min}}$ are the maximum and minimum attribute values in the case memory, respectively.

Figure 5. Distribution of similarity in case of point matching function

**ii) Interval matching function**

An interval matching function can be chosen for the numeric value when $f_i$ is in the form of a specific interval. A good example is a motor with revolution speed range of 0 to 2500 rpm. In Fig. 6, $R$ is the maximum range of attributes in the case memory and $r$ is the range of a specific case or $P$. In this case, the similarity can be obtained by simply dividing $r$ by $R$.

Figure 6. The interval matching function

By taking together all similarity measurements, the total similarity or distance will be calculated as follows:

$$
\text{dist}(c_i,c_j) = 1 - \sum_{i=1}^{k} \omega \times \left[ \text{sim}(F_i^P, F_i^A) + \text{sim}(B_i^P, B_i^A) + \text{sim}(S_i^P, S_i^A) \right]
$$

(4)

where, $\omega$ is a weight for function, behavior and structure, $\text{sim}(F_i^P, F_i^A)$ is the similarity of the functional attributes between cases $A$ and $B$, $\text{sim}(B_i^P, B_i^A)$ is the similarity of the behavioral attributes between cases $A$ and $B$, and $\text{sim}(S_i^P, S_i^A)$ is the similarity of structural attributes between cases $A$ and $B$.

In equation (4), higher similarity values map to smaller distances. This distance will be used for $k$-medoids clustering as well as case retrieval and case adaptation. The weighting, $\omega$, should be determined by the designer and product domain.

**E. Case retrieval and selection**

As described in section C, once we have clustered all cases, case retrieval can be accomplished easily. Namely, we can select the nearest cluster by comparing only $n$-medoids with $P$. In Fig. 7, cluster 1 ($C_1$) is selected as the nearest cluster to $P$ because the medoid of cluster 1 ($v_1$) has the shortest distance to $P$.

Figure 7. Structure of eco impact value and the contribution of structural attribute

**F. Case adaptation**

The environmental effect of a product is comprised of all impacts throughout the entire life cycle (raw material acquisition, manufacturing process, transportation, use and disposal etc., see [14]). The eco impact of each particular life stage can be divided into several impact categories. The structural attributes have strong influence on particular
impact categories, because manufacturing type, operation time and delivery mode, etc. are largely affected by structural attributes. The Fig. 9 describes these relations schematically and provides the relevant equations. The upstream process which is comprised of raw material acquisition and part manufacturing has \( l \)-impact category. Among them, for example, impact category value of ‘fossil fuels’ will be comprised of the impact of structural attributes such as length, diameter, thickness and surface area etc. Of course the exact value will be very complicated, but one exists. Therefore, we attempt to model these relationships from cases in the set \( C_p \) that have quite similar eco impact.

Based on above conceptual idea, we select some geometric properties that seems to have a greater impact as adaptation attributes, and we assume that the eco impact from these adaptation attributes will have similar tendencies within the set \( C_p' \). To figure out these tendencies, an impact category value is defined by the sum of each product of optimal \( k \)-set and adaptation attribute. Thus, we assume that the environmental effect of the upstream process can be described as:

\[
LS_{ups} = \sum_{i=1}^{n} \sum_{j=1}^{m} \left( k_j \times ad_{ij} \right)
\]

(5)

Fig. 10 shows the conceptual idea to estimate the eco impact of \( P \) using other cases in the new cluster set. In Fig. 10, the true \( l \)-th eco impact category value of case 3 is \( c_3, IC_l \), and the estimated value of case 3 by optimal \( k \)-set is \( c_3, elIC_l \). The difference between \( c_3, IC_l \) and \( c_3, elIC_l \) is our estimation error.

![Figure 10. Optimal k-set for \( C_p' \) and estimation result](image)

If we can find an optimal \( k \)-set that minimize the sum of errors between the true and estimated value of all cases in \( C_p' \), we can also estimate the eco impact value of \( P \) \( (P, elIC_p) \) by adapting the \( k \)-set to \( P \) because case 3, 6, 7 and \( P \) are all in \( C_p' \).

This procedure is defined as the following new optimization problem:

**New optimization problem of \( C_p' \)**

**Known:** Geometry attribute value of selected similar cases \((ad_{ij})\) and their LCA result \((c_3, s_{ij})\)

**Unknown:** LCA value of \( P, s_{ij} \)

**Find** optimal \( k \)-set satisfying following object function

\[
Obj \ Function = \min \left\{ \sum_{i=1}^{n} \left( c_3, IC_l - c_3, elIC_l \right) \right\}
\]

Finally, the complicated LCA procedure can be substituted by the above optimization problem.

### III. Case study

The effectiveness of the proposed CBR for LCA method is illustrated with a demonstrative example of a cross flow fan mounted on vehicle air purifier.

#### A. Outline of case study

We selected a cross flow fan in the vehicle air purifier as a new problem \( P \). Fig. 11 shows a photograph of the vehicle air purifier, fan and the abridged specifications. Our focus is only on the upstream process including raw material acquisition and part manufacturing process, not the downstream processes.

![Figure 11. Cross flow fan and specifications](image)

This cross flow fan is composed of 4 insert plates and 26 blades. We assume that the manufacturing process is largely divided into raw material cutting and the assembly process of plates and blades.

#### B. Case memory organization by \( k \)-medoids clustering

First, we collected one hundred varieties of fan to serve as cases in our case memory organization; see Fig. 12.

![Figure 12. Case memory classification by impeller profile](image)

The similarity between each case was calculated. The similarity between each case was calculated. The cases were then divided into five clusters as shown in Fig. 13.

![Figure 13. Case distribution in each cluster](image)

#### Classification of Impeller/Blades profile

By FBSE expressions, all cases in case memory were modeled. The similarity between each case was calculated. The cases were then divided into five clusters as shown in Fig. 13.
In Fig. 13, only the distance between a medoid and the other cases in each cluster has meaning to clustering, and relative distance between each cluster was not represented in Fig. 13.

C. Case retrieval and selection for adaptation

In order to select cases for adaptation, the distance among the five medoids of each cluster and \( P \) were calculated. As shown in Table I, cluster 3 was the closest cluster to \( P \). We expect the cases in cluster 3 to exhibit environmental effects that are more similar to \( P \) than the cases in the other clusters.

<table>
<thead>
<tr>
<th>Table I. Distance on ( P ) of each cluster medoid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster No.</td>
</tr>
<tr>
<td>Medoid No.</td>
</tr>
<tr>
<td>Distance</td>
</tr>
</tbody>
</table>

As shown in Fig. 14 and Fig. 15, the proposed algorithm can estimate the eco impact value within a few percent error. This is quite a reasonable from the adaptation cases.

IV. CONCLUDING REMARKS

For use in eco product development, we develop an approximate LCA method using CBR for a rapid and quantitative environmental evaluation. The results of our study have confirmed the importance of the following important concepts:

- FBSE expression for CBR
- Similarity measurement for case indexing and retrieval
- Eco impact value estimation algorithm based on geometric attribute characteristics

However, if there are no cases similar to \( P \) in the case memory or, if the case memory does not have a sufficient number of cases, then the proposed algorithm may not predict the eco values as accurately. These limitations are innate properties of CBR. An extended algorithm capable of addressing the shortcomings is currently under investigation.

REFERENCES