International Journal of Computational Intelligence Systems, Vol. 1, No. 1 (January, 2008) 24–49

ASSESSMENT OF STRATEGIC R&D PROJECTS FOR CAR MANUFACTURERS BASED ON THE EVIDENTIAL REASONING APPROACH

Xin-Bao LIU[1], Mi ZHOU[1], Jian-Bo YANG[2], Shan-Lin YANG[1,*]

[1] *School of management, Hefei University of Technology, 230009 Hefei, Anhui , P.R. China*

[2] *Manchester Business School, The University of Manchester, Manchester M15 6PB, UK*

http://www.hfut.edu.cn

Received September 21th, 2007
Revised September 27th, 2007

Assessment of strategic R&D projects is in essence a multiple-attribute decision analysis (MADA) problem. In such problems, qualitative information with subjective judgments of ambiguity is often provided by people together with quantitative data that may be imprecise or incomplete. A few approaches can be used to deal with such quantitative and qualitative MADA problems under uncertainty, such as the evidential reasoning (ER) approach that has its own unique features. In this paper, the ER approach is applied to the assessment of strategic R&D projects for a car manufacturer, which is characterized by many qualitative factors that may be imprecise or fuzzy. The ER approach is well-suited for dealing with such problems and can generate comprehensive distributed assessments for different projects. The group analytic hierarchy process (GAHP) method is applied to calculate the weights of attributes in the E-R assessment process, where a group of people from the company were involved. We also provide a new algorithm for the comparison of two alternatives under utility interval. Our research that has been undertaken for the car manufacturer has contributed to the improvement of the quality and efficiency of its strategic R&D projects. The research has also helped the personnel of the company better understand the benefits of using scientific methods for systematic project assessment.

Keywords: Strategic R&D project assessment; Evidential reasoning; Multiple-attribute decision analysis; Qualitative and quantitative information; Utility; GAHP

1. Introduction

Research and development (simply R&D) project assessment is concerned with the general evaluation and checkup of R&D projects based on a scientific evaluation system with appropriate criteria. Using a reliable and rational evaluation system to assess R&D projects is very important for a company to enhance the effectiveness of assessment and improve its product quality, which ultimately leads to the improvement of its overall performance.

Theoretically, R&D project assessment could be categorized into the domain of project evaluation,[30,31,32,33,34] which includes systematic analysis and taking actions in the process of project decision making and implementation. One of the important tasks in planning large and advanced R&D projects is to minimize innate uncertainties and ambiguities in the management of projects.[35] In the car manufacturing industry, the investment of a strategic R&D project is always concerned with a great deal of money and human resources and is re-

*Corresponding authors. Address: School of management, Hefei University of Technology, 230009 Hefei, Anhui, P.R. China (S.L. Yang).
Tel.: +86 551-2901008.[1,2] Tel.: +86 551-2904989.[3] Tel.: +44 161 200 3427.
E-mail: lxinbao@mail.hf.ah.cn; zhoumi83@163.com; jian-bo.yang@manchester.ac.uk; slyang@mail.hf.ah.cn

lated to the investors' long time interests and future markets. Failure in design or production can lead to deadly consequences to a car manufacturer. If the launch of a new type of car is delayed, its market opportunities may be lost. It is therefore vital to assess R&D projects to avoid unnecessary waste of money and time.

Project management is a series of systematic management for the achievement of project goals under the constrained resources.[36,37] The modern project management is generally considered to be started in 1940s. There are many approaches for project management, the kernel of which is the process of establishing shadow price of a project, which covers the consideration of social effects and responsibilities from the project, apart from the optimization of profit only. This is evidently distinctive from earlier research.

In a strategic R&D project evaluation problem, various types of attributes need to be taken into account, which may be quantitative, measured by numerical values with certain units, or qualitative, assessed using subjective judgments with uncertainties. In such a multiple attribute decision analysis (simply MADA) problem, subjective judgments are often provided by a group of assessors because an individual may be incapable of providing reliable judgments due to the lack of information and/or experiences. In a traditional MADA problem, several quantified evaluation grades may be defined for assessing an attribute, and a numerical value associated with the assessed grades could then be used to evaluate an alternative on an attribute. There are a number of approaches that can be used to deal with the traditional MADA problem, for example AHP,[8,9] TOPSIS,[4] ELECTRE-I (Roy, 1971),[27] ELECTRE-II (Roy, 1975),[28] PROMETHEE (Brans, 1984),[29] and so on. These approaches are not suitable to deal with problems with subjective judgments which could not be quantified appropriately. The problem arises as to how to assess qualitative attributes with imprecise information in R&D project evaluation. Over the last 20 years, a lot of research has been conducted to develop methods for dealing with uncertain information.[1,2,5,16,18]

The evidential reasoning (ER) approach was therefore introduced in 1990s [19,20,22,24] based on the Dempster–Shafer (D-S) theory[3,7] and decision theory. This approach is well-suited to addressing uncertain MADA problems with qualitative attributes in strategic R&D project assessment in a rational way since the concept of qualitative evaluation grades is imbedded in the assessment process and qualitative evaluation grades do not have to be quantified in the ER-based assessment aggregation. The unique features of the ER approach include its ability to represent incomplete and vague subjective judgments and its convenience for the combination of attributes, which will be discussed in detail in this paper. In resent years, the ER approach has been applied in a number of areas, for example the environmental impact assessment,[15] organizational self-assessment,[23] pre-qualifying construction contractors,[11] general cargo ship design,[10] motorcycle assessment,[19] and marine system safety analysis and synthesis.[12]

In this paper, the ER approach will be applied for the assessment of R&D projects for car manufacturers for the first time, and weights in the assessment model are acquired through on-site investigations in a car manufacturer. In Section 2, the basics of the Dempster–Shafer's theory of evidence will be described. Section 3 is intended to discuss the assessment framework of the ER approach in the context of R&D project assessment. In Section 4, the recursive ER algorithm will be introduced. Section 5 is devoted to the application of the ER approach to the assessment of the R&D projects for a car manufacturer, and sensitivity analysis of weights from different departments is conducted. The paper is concluded in Section 6.

2. Basics of Dempster–Shafer's theory of evidence

Dempster–Shafer's evidence theory is well suited for handling incomplete information. It is introduced by Dempster in Ref. [3] and refined by Shafer,[7] so the theory of evidence is called the D-S theory of evidence. In the D-S theory of evidence, a basic hypothesis (proposition) is denoted by H_n. All hypotheses together constitute a set $\theta = \{\mathrm{H}_1, \mathrm{H}_2, \ldots, \mathrm{H}_N\}$, which is called the frame of discernment (sample space). The hypotheses in θ are collectively exhaustive and mutually exclusive and the elements in θ could be enumerated by 2^θ which is the power set of θ, consisting of all the subsets of θ. Suppose there are two basic hypotheses A and B in θ, then the frame of discernment is $\theta = \{\mathrm{A}, \mathrm{B}\}$ and $2^\theta = \{\varnothing, \{\mathrm{A}\}, \{\mathrm{B}\}, \{\mathrm{A}, \mathrm{B}\}\}$.

Let $m(\mathrm{A})$ denote the basic probability assignment (mass) to the subset A, which measures the extent to which the evidence supports A. It is a number between $[0,1]$ satisfying the following two equations

$$\sum_{\mathrm{A}\subseteq\theta} m(\mathrm{A}) = 1, \quad 0 \leqslant m(\mathrm{A}) \leqslant 1, \quad \forall \mathrm{A} \subset \theta \tag{1}$$

$$m(\varnothing) = 0 \tag{2}$$

A is called a focal element if it satisfies $m(\mathrm{A}) > 0$, $\mathrm{A} \subseteq \theta$, and all of the focal elements together are the core of θ. $m(\mathrm{A})$ expresses the probability mass exactly assigned to A but not to any subset of A.

The assigned probability to θ which is denoted by $m(\theta)$ is the measurement of the degree of ignorance. It is assumed to be the negation of the hypothesis A if A is the only focal element. For example, if $m(\mathrm{A}) = r$, $\mathrm{A} \subseteq \theta$, and all other subsets of θ are not assigned any probability mass, then we will get $m(\theta) = 1 - r$.

Belief function is another important concept associated with the evidence theory which is defined as

$$\mathrm{Bel}(\mathrm{A}) = \sum_{\mathrm{B}\subseteq\mathrm{A}} m(\mathrm{B}) \quad (\forall \mathrm{A} \subset \theta) \tag{3}$$

It reflects the exact support to the hypothesis A and is a function $\mathrm{Bel} : 2^\theta \to [0,1]$. $\mathrm{Bel}(\mathrm{A})$ is the probability assigned A to considering all the premises of A. There are several other functions associated with the evidence theory such as the plausibility function, commonality function and so on. Each of them reflects the probability based on the basic probability assignment (number) from different points of view under the frame of discernment.

The kernel of the D-S theory of evidence is the combination rule which could be used for the aggregation of different sources of evidence. Suppose there are n pieces of evidence in θ, and they each provide a basic probability assignment to a subset A of θ, i.e. $m_1, m_2, \ldots, m_n$. The evidence combination rule is defined as follows:

$$\begin{aligned} K &= \left(1 - \sum_{\substack{\mathrm{A}_1,\mathrm{A}_2,\ldots,\mathrm{A}_n\subset\theta \\ \mathrm{A}_1\cap\mathrm{A}_2\cap\cdots\cap\mathrm{A}_n=\varnothing}} m_1(\mathrm{A}_1) m_2(\mathrm{A}_2)\cdots m_n(\mathrm{A}_n)\right)^{-1} \\ &= \left(\sum_{\substack{\mathrm{A}_1,\ldots,\mathrm{A}_n\subset\theta \\ \mathrm{A}_1\cap\cdots\cap\mathrm{A}_n\neq\varnothing}} m_1(\mathrm{A}_1) m_2(\mathrm{A}_2)\cdots m_n(\mathrm{A}_n)\right)^{-1} \end{aligned} \tag{4}$$

$$m(\mathrm{A}) = \begin{cases} 0, & \mathrm{A} = \varnothing \\ K \cdot \sum\limits_{\substack{\mathrm{A}_1,\ldots,\mathrm{A}_n\subset\theta \\ \mathrm{A}_1\cap\cdots\cap\mathrm{A}_n=\mathrm{A}}} m_1(\mathrm{A}_1)\cdots m_n(\mathrm{A}_n), & \mathrm{A} \neq \varnothing \end{cases} \tag{5}$$

where K is called the normalization factor and $K > 1$. When $\mathrm{A}_1 \cap \mathrm{A}_2 \cap \cdots \cap \mathrm{A}_n = \varnothing$, basic probabilities are assigned to n pieces of inconsistent evidence that lead to the conflict among them. The larger K is, the greater the conflict is. So K is called the degree of conflict that reflects the conflict between n pieces of evidence. In the combination rule, it is assumed that the information sources are independent and the process of calculation is called the orthogonal sum.

Dempster's rule of combination satisfies commutativity and associativity of multiplication. As such, it ensures that the combination results remain the same regardless of the order in which the n pieces of evidence are aggregated.

3. ER-based general framework for assessment of R&D projects

3.1. R&D project assessment based on the ER framework

A hybrid MADA problem is comprised of multiple technical and economical attributes (criteria), which may be either quantitative or qualitative. Suppose there are t alternatives to be assessed on k quantitative factors and h qualitative factors. In such a hybrid MADA decision problem, subjective judgments with uncertainty may be provided by assessors and aggregated by evidence based reasoning.

In R&D project assessment, for example, "theoretical value and level of innovation" is regarded to be good if "theoretical value", "innovation" and "individual design" are all assessed to be good. However, it is rarely the case that assessments could always be as precise as this. For example, assessors may select an evaluation grade or more grades with different probability assignments. In the real assessment of R&D projects, for example, "theoretical value of project" associated with "quality of production" may be stated as follows. "The theoretical value of project for heavy trailer is evaluated to be best with a belief degree of 0.42, to be good with a belief degree of 0.50, and to be poor with a belief degree of 0.08". The above statement could be represented by the following expectation:

$$\begin{aligned} &\mathrm{S}(\text{Theoretical value}) \\ &\quad = \{(\text{Best}, 0.42), (\text{Good}, 0.50), (\text{Poor}, 0.08)\}. \end{aligned}$$

where $\mathrm{S}(\text{Theoretical value})$ stands for the state of the heavy trailer's "theoretical value".

Note that the total belief degree for the statement of "theoretical value of project" sums to 1.0, which means that the information provided by experts is complete. If the belief degree to each evaluation grades do not sum to one, it means that the information provided is incomplete.

From the above statements, a set of evaluation grades for the assessment of a factor on an alternative is defined as follows.

$$\mathrm{H} = \{\mathrm{H}_1, \mathrm{H}_2, \ldots, \mathrm{H}_N\} \tag{6}$$

where H_n $(n = 1, 2, \ldots, N)$ each denotes an evaluation grade. They are collectively exhaustive and mutually exclusive.[22] H_1 represents the least preferred evaluation grade and H_N represents the most preferred evaluation grade. H_{n+1} is supposed to be preferred to H_n. In the above R&D project assessment, for example, there are 5 evaluation grades associated with the factor "theoretical value", which are defined as

$$\mathrm{H} = \{\text{Worst}, \text{Poor}, \text{Average}, \text{Good}, \text{Best}\} \tag{7}$$

For illustration purpose, best is represented by A, good by B, average by C, poor by D, and worst by E. Then

$$\mathrm{H}^{\text{general}} = \{\mathrm{E}, \mathrm{D}, \mathrm{C}, \mathrm{B}, \mathrm{A}\} \tag{8}$$

In R&D project assessment, a factor may have its unique set of evaluation grades which is different from other factors to facilitate raw data collection.[22] Then, it is essential to transform various sets of evaluation standards associated with both quantitative and qualitative attributes to a unified set using the assessor's knowledge without changing the features of incomplete assessments.[22] The process will be introduced in the following subsection.

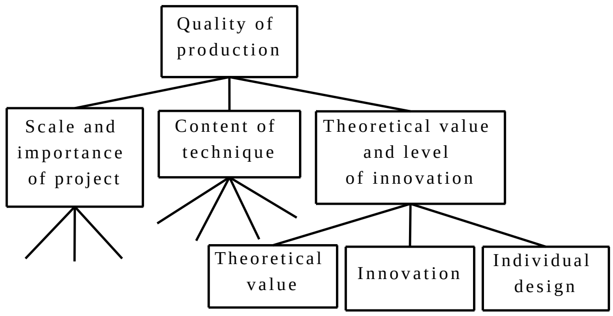

Fig. 1. The hierarchical structure for "quality of production" in R&D project assessment

Hierarchical analysis is common in MADA problems. In the assessment of motorcycle,[19] for example, there are two hierarchies of attributes. It is extended to four levels of attributes in Ref. [20] for more precise evaluation. In this paper, "quality of production" is an abstract concept that could not be judged directly, and therefore is decomposed into three detailed concepts: "scale and importance", "technique" and "theoretical value and level of innovation". If it is still too abstract to assess them directly, then they should be decomposed into more detailed factors until they could be evaluated directly. In our research, "quality of production" is a hierarchical structure of three levels shown in Fig. 1.

3.2. Evaluation model

In the above R&D project assessment framework, "Theoretical value", "Innovation" and "Individual design" are referred to as three basic factors associated with its upper attribute "theoretical value and level of innovation" in the second level of the hierarchy. In general, suppose there are L attributes in the top level which are defined as

$$E = \{E_i,\, i = 1,2,\ldots,L\} \qquad (9)$$

It contains a complete set of factors for the evaluation of the general assessment. Suppose there are j_i attributes associated with E_i, which is denoted as

$$E_i = \{E_{ij},\, j = 1,2,\ldots,j_i\} \qquad (10)$$

Taking "quality of production", E_1 for example, three sets of attributes in the second level for assessing it are defined as follows:

$$E_1 = \{E_{11}, E_{12}, E_{13}\} = \{\text{scale and importance of project, content of technique, Theoretical value and level of innovation}\},\ j_1 = 3$$

Suppose there are s_{ij} basic attributes for assessing E_{ij}, which could be denoted as follows:

$$E_{ij} = \{e_{ij}^k,\, k = 1,2,\ldots,s_{ij}\} \qquad (11)$$

$$E_{11} = \{e_{11}^1, e_{11}^2, e_{11}^3\} = \{\text{workload, origin of person, importance of project}\},\ s_{11} = 3$$
$$E_{12} = \{e_{12}^1, e_{12}^2, e_{12}^3, e_{12}^4\} = \{\text{complexity of critical technique, ratio of quality/price, reliability, efficiency}\},\ s_{12} = 4$$
$$E_{13} = \{e_{13}^1, e_{13}^2, e_{13}^3\} = \{\text{theoretical value, innovation, individual design}\},\ s_{13} = 3$$

The state of a basic factor e_{ij}^k evaluated for a R&D project a_t to a grade H_n may then be described as the following expectation:

$$S(e_{ij}^k(a_t)) = \left\{(\mathrm{H}_n, \beta_{n,ij}^k(a_t)),\ n = 1,2,\ldots,N\right\} \quad (i = 1,\ldots,L;\ j = 1,\ldots,j_i;\ k = 1,\ldots,s_{ij};\ t = 1,\ldots,M) \qquad (12)$$

where $\beta_{n,ij}^k(a_t)$ expresses the intensity to which the state of a single factor e_{ij}^k at a_t is assessed to an evaluation grade H_n. In Ref. [19], $\beta_{n,ij}^k(a_t)$ is assumed to satisfy a rationality assumption in which $0 \leqslant \beta_{n,ij}^k(a_t) \leqslant 1$ is commonly satisfied whereas the other two assumptions are acquired only by experience. The state of e_{ij}^k at a project could be evaluated to any evaluation grades [20] instead of two adjacent grades.[19] In the statement, $\sum_{n=1}^{N} \beta_{n,ij}^k(a_t) \leqslant 1$ is a basic condition for $\beta_{n,ij}^k(a_t)$. If $\beta_{n,ij}^k(a_t)$ $(n = 1,2,\ldots,N)$ sums to unity, the information provided by DM is said to be complete. If $\sum_{n=1}^{N} \beta_{n,ij}^k(a_t) < 1$, the judgment is incomplete. Based on the above expectation, a belief decision matrix may be modeled as[19,22]:

$$D_g = \left(S(e_{ij}^k(a_t))\right)_{S\times M} \qquad (13)$$

S is the total number of basic attributes. To quantify the evaluation grades of the basic or general attributes, a function must be defined for H_n. In Ref. [19] and [20], the concept of preference degree $p(\mathrm{H}_n)$ is introduced and a utility function $u(\mathrm{H}_n)$ is used in Ref. [22,24,15] and [11]. $p(\mathrm{H}_n)$ takes the value in the close interval $[-1, 1]$, and $u(\mathrm{H}_n)$ is estimated from zero to one. There are a lot of utility estimation methods. For example, three approaches to estimate utilities on a quantitative attribute are discussed in Ref. [22]. In this paper, we assume that the utility of evaluation grades are evenly distributed as follows:

$$u(A) = 1, \quad u(B) = 0.75, \quad u(C) = 0.5, \quad u(D) = 0.25, \quad u(E) = 0. \qquad (14)$$

4. Introduction to the E-R algorithm for assessment aggregation

4.1. The recursive ER algorithm

In the above R&D project assessment framework, if "scale and importance", "technique" and "theoretical value and level of innovation" are all assessed to be good, then the overall assessment to "quality of production" will be good as well. But it is not always the case. So the question is how to aggregate the assessments on different factors to generate an overall assessment.

Based on the D-S theory and an assessment framework as briefly described above, the ER algorithm is developed in Ref. [19] and improved in several other papers.[15,20,22,24] The recursive ER algorithm is briefly described as follows. Suppose there are L basic attributes in assessing the general attribute E, where $E = \{e_1, \ldots, e_L\}$. A recursive ER algorithm is given as follows:

$$\overline{m}_{H,i} = 1 - \omega_i, \qquad (i = 1,2,\ldots,L) \tag{15}$$

$$\widetilde{m}_{H,i} = \omega_i \left(1 - \sum_{n=1}^{N} \beta_{n,i}\right) \text{with } m_{H,i} = \overline{m}_{H,i} + \widetilde{m}_{H,i}$$
$$\text{and } \sum_{i=1}^{L} \omega_i = 1 \qquad (i = 1,2,\ldots,L) \tag{16}$$

$$E_{I(i)} = \{e_1, \ldots, e_i\} \qquad (i = 1,2,\ldots,L) \tag{17}$$

$$\mathrm{H}_n : m_{I(i+1)} = K_{I(i+1)} \left[m_{n,I(i)} m_{n,i+1} + m_{H,I(i)} m_{n,i+1} + m_{n,I(i)} m_{H,i+1} \right] \tag{18}$$

$$m_{H,I(i)} = \widetilde{m}_{H,I(i)} + \overline{m}_{H,I(i)}, \quad (n = 1,2,\ldots,N) \tag{19}$$

$$\mathrm{H} : \quad \widetilde{m}_{H,I(i+1)} = K_{I(i+1)} \left[\widetilde{m}_{H,I(i)} \widetilde{m}_{H,i+1} + \overline{m}_{H,I(i)} \widetilde{m}_{H,i+1} + \widetilde{m}_{H,I(i)} \overline{m}_{H,i+1} \right] \tag{20}$$

$$\mathrm{H} : \quad \overline{m}_{H,I(i+1)} = K_{I(i+1)} \left[\overline{m}_{H,I(i)} \overline{m}_{H,i+1} \right] \tag{21}$$

$$K_{I(i+1)} = \left[1 - \sum_{s=1}^{N} \sum_{\substack{j=1 \\ j \neq s}}^{N} m_{s,I(i)} m_{j,i+1} \right]^{-1}, \quad (i = 1,2,\ldots,L-1)$$

ω_i is the relative importance of factor e_i, and $m_{n,i}$ represents the degree to which e_i supports a hypothesis that the state of the attribute E at an alternative a_t is assessed to H_n.[3] $E_{I(i)}$ represents the first i factors associated with their upper level attribute, and $m_{n,I(i+1)}$ is the combined probability assignment to H_n generated by assessing $E_{I(i)}$. $m_{n,I(1)} = m_{n,1}$ ($n = 1,2,\ldots,N$) and $m_{H,I(1)} = m_{H,1}$. From the formulae, it is clear that $L-1$ rounds of calculation need to be conducted for the generation of the overall assessment $m_{n,I(L)}$ and $m_{H_n,I(L)}$ by combining the L basic factors. Let β_n be the belief degree to which the general attribute E is assessed to the grade H_n and β_H be the belief degree unassigned to any individual grade. Then

$$\mathrm{H}_n : \beta_n = \frac{\widetilde{m}_{H,I(L)}}{1 - \overline{m}_{H,I(L)}} \qquad (n = 1,2,\ldots,N) \tag{22}$$

$$\mathrm{H} : \beta_H = \frac{m_{n,I(L)}}{1 - \overline{m}_{H,I(L)}} \tag{23}$$

In the above formulae, $\widetilde{m}_{H,i}$ represents the remaining probability mass initially unassigned to any individual grades, caused by the incompleteness of the assessment on the basic R&D factor i. $\overline{m}_{H,i}$ is caused by the relative importance of R&D factors. In other words, it represents how the other factors could contribute to evaluating the general attribute (E).

After the aggregation of L basic factors, a distributed assessment for a R&D project a_t on the general attribute E can then be presented as follows:

$$S(y(a_t)) = \{(\mathrm{H}_n, \beta_n(a_t)),\ n = 1, 2, \ldots, N\} \tag{24}$$

4.2. Ranking based on expected utility and utility interval

In a R&D project assessment problem, it may not be straightforward to arrive at the ranking of projects only based upon the distributed assessments as shown above. For example, a project may be evaluated to be "average" to a very large extent, whereas another project may be evaluated to "poor" and "good" each to similar large degrees. Therefore, it will be difficult to differentiate these two projects based only on the distributed assessments of the two projects. In such circumstances, it is necessary to generate a utility value or a numerical score for a project. For this purpose, a utility must be estimated for each evaluation grade.

Suppose $u(\mathrm{H}_n)$ ($n = 1,2,\ldots,N$) denotes the utility of an evaluation grade H_n and $u(\mathrm{H}_{n+1})$ is assumed to be larger than $u(\mathrm{H}_n)$ if H_{n+1} is preferred to H_n. The most preferred grade H_N is supposed to be of the highest value, whereas the least preferred grade H_1 has the lowest value. The expected utility of the state of attribute y_k on project a_t may then be calculated by the following expectation.

$$u(S(y(a_l))) = \sum_{n=1}^{N} u(\mathrm{H}_n) \beta_n(a_l) \tag{25}$$

As mentioned before, $\beta_{\rm H}$ denotes the degree of belief unassigned to any individual grades. So, if the assessment of attribute y_k is incomplete, $\beta_{\rm H}$ will be positive. Thus, the belief degree $\beta_n(a_t)$ denotes the lower bound of the likelihood that a R&D project is assessed to H_n, while $\beta_n(a_l)+\beta_{\rm H}(a_l)$ represents the upper bound of the likelihood.[24] The degree that y_k may be assessed to H_n is anything in the interval $[\beta_n(a), \beta_n(a_l)+\beta_{\rm H}(a_l)]$. Based on the belief interval, maximum, minimum and average utilities of a project could be measured by the following formulae:

$$u_{\max}(a_l) = u(S(y(a_l))) + \beta_{\rm H}(a_l)u(\mathrm{H}_N) \tag{26}$$

$$u_{\min}(a_l) = u(S(y(a_l))) + \beta_{\rm H}(a_l)u(\mathrm{H}_1) \tag{27}$$

$$\begin{aligned} u_{\rm ave}(a_l) &= \frac{u_{\max}(a_l)+u_{\min}(a_l)}{2} \\ &= u(S(y(a_l))) + \beta_{\rm H}(a_l)\frac{u(\mathrm{H}_N)+u(\mathrm{H}_1)}{2} \end{aligned} \tag{28}$$

From the above equations, we could see that $u_{\min}(a_1) = u(S(y(a_1)))$ if $u(\mathrm{H}_1) = 0$, and $u_{\min}(a_1) = u_{\max}(a_1) = u_{\rm ave}(a_1) = u(S(y(a_1)))$ if the original information provided by the assessors is complete. The difference $(u_{\max}(a_1) - u_{\min}(a_1))$ reflects the degree of the incompleteness in the assessment. It is a linear increasing function of $\beta_{\rm H}(a_1)$. Based on the utility interval, an equation for ranking two R&D projects is provided as follows:

$$p(a_l > a_k) = \min\{\max(\Lambda, 0), 1\} \tag{29}$$

where

$$\Lambda = \frac{u_{\max}(a_l) - u_{\min}(a_k)}{[u_{\max}(a_l) - u_{\min}(a_k)] + [u_{\max}(a_l) - u_{\min}(a_k)]}$$

a_l is regarded to be superior to a_k to a degree of $p(a_l > a_k)$ if $p(a_l > a_k) > 0.5$; if $p(a_l > a_k) < 0.5$, then a_l is supposed to be inferior to a_k to a degree of $1 - p(a_l > a_k)$; if $p(a_l > a_k) = 0.5$, then a_l is regarded to be indifferent to a_k.

5. Application: checkout and assessment system of strategic R&D projects

In this section, the evidential reasoning approach is applied to analyze the performance of several R&D projects for a car manufacturer in China. The research was conducted in close collaboration with the leaders in the marketing department, manufacturing department, and human resource department of the company, whose name is not mentioned here to protect its business interests. The meetings were also held with the technical and non-technical staff in both the technique and non-technique centers of the company. The company has established a preliminary system of checkout and assessment. However, the evaluation of projects in the system is purely based on subjective analysis, some basic attributes are only roughly described, and some important attributes have not been established yet. Thus, there is a need to improve the current checkout and assessment system and provide a more complete set of criteria and a more scientific and reliable evaluation system. This section is devoted to the development of such a system.

5.1. *Description of the "strategic R&D project assessment" problem*

(1) Connotation about assessment of R&D projects

Checkout and assessment of a project refers to the systematic and objective analysis of the goal of the project, the process of the project implementation, and its benefits and influence at the end of the project. It is assumed that the following three objectives should be achieved through the R&D project assessment process. 1) People concerned with a R&D project should agree on what have been achieved from the project, 2) experiences in conducting the project should be collected, and 3) the quality of decision making and investment efficiency should be improved.

(2) Effect from assessment of R&D projects

The effect could be expressed from four aspects. Firstly, the assessment is an appropriate way for gathering experiences. Secondly, it is a tool for improving the quality of project decision making in organization. Thirdly, it is a requirement for the continual development of projects. Finally, the evaluation process is supposed to facilitate information feedback in time, regulate the contents and approaches for decision making, improve project decision making process and increase return from investment.

(3) The contents about assessment of R&D projects

The checkout and assessment of a project is analyzed by assessing the necessity of carrying out the project, the quality of production and its potential in market, advancement in technology, and feasibility assessment with respect to finance and economy. The detailed contents of the assessment system will be presented in the next subsection.

5.2. *Identification of factors for "checkout and assessment system of strategic R&D projects"*

5.2.1. General attributes

(1) Quality of production

Quality of production contains two aspects. One aspect is concerned with qualitative factors which are characterized by the contents of techniques used in projects. Quality is defined as the overall satisfaction generated from the stated characteristics of a product, which means how the consumer's demand could be satisfied by the features of a product. Production quality can be assessed from the following four areas. The first area is related to the performance of a product for use, which means the possessed technical characteristics with respect to its usage. The second one is characterized by reliability, or the ability of a product to accomplish the intended functions over the given time and conditions, which is generally measured by invalidation rate, average time of no-malfunction, and so on. The third one is related to security in circulation, usage and operation of a product. The last one is about economy. Economy is the cost spent on designing, manufacturing and using a product. Based on these characteristics and for the assessment of R&D process, we use the following factors to measure the contents of technique:

- complexity of critical techniques in project
- ratio between quality and price
- reliability of product
- economy

The other aspect about production quality is related to quantitative factors to measure the scale and importance of a project, where the scale of a project is evaluated by the workload (concerned with working days) and origin of persons (concerned with departments).

For the checkout and assessment of a strategic R&D project, it is necessary to consider the theoretical value and level of innovation in the evaluation process apart from the contents of technique. For more detailed assessment, level of innovation is split into the extent of innovation and the ratio of innovation in individual designs.

(2) Process control

Not much attention is paid to process control in the original project evaluation system used by the company. In our investigation, we recognized that the research personnel and project directors of the company did care about process control. In this paper, process control is defined as quality control, time control and investment control.

(3) Added value

During the investigation, it is recognized that achievements on personnel, project process and product are made after the completion of a project, whilst project process is regarded to be the most valuable factor. Compared with the product which is the direct result after a project, it seems to be more significant for a company to have built up an excellent project group and the management methods of the whole project process because they can help improve management decision making in future. Therefore, in addition to evaluating product quality in the assessment of strategic R&D projects, the quality of the whole project process needs to be assessed. In this process, the methods used and experiences gained need to be documented by the project group, and the continuity of techniques associated with a project should be covered as well.

The added value from a project, which includes experienced project group and technique continuity, is very important because it can help improve project management and generate more benefits in the development of new projects in future. An experienced project group is evaluated by the following three parts.

- Documents of rules and regulations established by the project group
- The routine operational management documents
- Management documents about R&D process of product

The routine operational management documents and management documents about R&D process of product are defined as follows. The project management is handled in two ways in the present R&D system. First, the management is about routine project operation. Product development consists of five steps: planning, development, business development, trial-manufacturing and evaluation. The personnel should be told what to do and how to do at each step. Any problems should not be allowed to be unresolved till the next step. The other one is regarded as the management about R&D process of product. People involved in a project should be made aware what they need to do in each step of a project and how to solve problems in the development of a project. They also need to decide how many documents should be created to ensure the smooth development of a project system.

Compared with the system of rules and regulations, group coherence and communication are also important for the R&D assessment, but they are to a large extent subjective and difficult for explicit quantification.

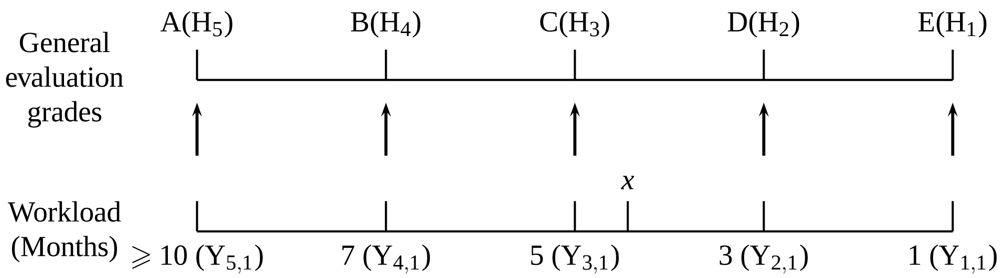

Fig. 2. Transformation Rules for Workload

5.2.2. *The evaluation grades for each basic quantitative or qualitative attribute*

17 attributes are selected for the evaluation of the R&D projects for the car manufacturer, as shown in Table 2. In the previous sections of this paper, the general attribute structure was studied. In this subsection, a detailed structure of the evaluation grades on each attribute is investigated.

Quality of production

a. Workload

Workload is a quantitative attribute associated with product scale and importance. It is assumed to be the worst if a project is not complex and is intended to be finished within only one month, whereas a project is assessed to be the best if it is complex and is planned for 10 months or more. The evaluation grade is a continuous linear function in months. If a product is intended to be completed between 1 and 10 months, the evaluation grades should be transformed to the general grades by the rule-based information transformation techniques as shown in Fig. 2.[22]

Figure 2 shows that if a project consists of the workload of 10 or more, it would be evaluated to grade A. It means that the utility of 10 months is equivalent to the utility of A in the general evaluation grade. Similarly, the utility of 7 months is equivalent to the utility of B, 5 months to C, 3 months to D, and 1 month to E. So according to the principle of utility equivalence,[22] we have

$$u(Y_{n,i}) = u(\mathrm{H}_n) \;\; (n = 1, 2, \ldots, N) \text{ with } N = 5 \quad (30)$$

If the workload for a project is not one of the above five numbers, it should be transformed to the general grades by rules. Let $Y_{1,i}, \ldots, Y_{N,i}$ be N crisp assessment grades for quantitative attribute e_i. If the attribute value y_i lies between the two adjacent assessment grades $Y_{n,i}$ and $Y_{n+1,i}$, let $\beta_{n,i}$ and $\beta_{n+1,i}$ be the belief degrees to which y_i is assessed to these two grades. Then we have

$$\beta_{n,i} = \frac{Y_{n+1,i} - y_i}{Y_{n+1,i} - Y_{n,i}}, \quad \beta_{n+1,i} = \frac{y_i - Y_{n,i}}{Y_{n+1,i} - Y_{n,i}} \quad (31)$$

Thus, the quantitative attribute value y_i could be represented as follows:

$$S(y_i) = \left\{ (\mathrm{H}_n, \beta_{n,i}); (\mathrm{H}_{n+1}, \beta_{n+1,i}) \right\} \quad (32)$$

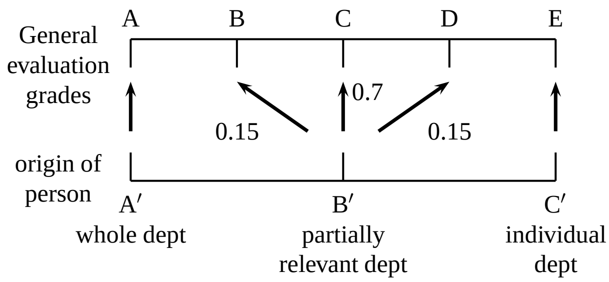

Fig. 3. Transformation Rules for Origin of Person

b. Origin of person

A′ - All departments
B′ - Some relevant departments
C′ - Single department

Here, three evaluation grades are defined for the assessment of "origin of person". If a project involves only one department, it would be marked C′, whereas if all departments join in, it would be assessed to the best grade A′. If a project includes some of the departments, it would be marked B′. These three evaluation grades for origin of person are transformed to the general 5 grades using the rule based information transformation techniques, as shown in Fig. 3.

The transformation matrix for origin of person is then given by

$$\overline{\mathrm{A}}\begin{pmatrix}\text{origin of}\\ \text{person}\end{pmatrix} = \begin{array}{c|ccccc} & \mathrm{A} & \mathrm{B} & \mathrm{C} & \mathrm{D} & \mathrm{E} \\ \mathrm{A}' & 1 & 0 & 0 & 0 & 0 \\ \mathrm{B}' & 0 & 0.15 & 0.7 & 0.15 & 0 \\ \mathrm{C}' & 0 & 0 & 0 & 0 & 1 \end{array} \quad (33)$$

It is assumed from Fig. 3, that A′ is equivalent to the general grade A, C′ to E, and B′ to a combination of 0.15B, 0.7C and 0.15D. In other words, $u(\mathrm{A}') = u(\mathrm{A})$, $u(\mathrm{C}') = u(\mathrm{E})$ and $u(\mathrm{B}') = 0.15u(\mathrm{B}) + 0.7u(\mathrm{C}) + 0.15u(\mathrm{D})$.

c. Importance of project

A′ - Very important
B′ - Important
C′ - Relatively important
D′ - Indifferent

A product is considered to be very important if this product can fill a gap in the product range of the company. If a product can fill a gap in a platform of product for the company, it should be considered to be important. If some improvement is made by developing a product, it should be considered to be relatively important, whereas it would be thought to be indifferent if no obvious improvement is made by developing a product. Figure 4 shows the relationship between the evaluation grades of this factor and the general evaluation grades.

The best grade A′ in this factor is considered to be equivalent to the general grade A, while D′ is equivalent to the general grade E. Thus, $u(\mathrm{A}') = u(\mathrm{A})$, $u(\mathrm{D}') = u(\mathrm{E})$. B′ is assigned to B and C simultaneously with the degrees of α_1 and β_1, and C′ to C and D with the degrees of α_2 and β_2. $u(\mathrm{B}')$ and $u(\mathrm{C}')$ could then be calculated as:

$$u(\mathrm{B}') = \alpha_1 u(\mathrm{B}) + \beta_1 u(\mathrm{C}) \quad (34)$$

$$u(\mathrm{C}') = \alpha_2 u(\mathrm{C}) + \beta_2 u(\mathrm{D}) \quad (35)$$

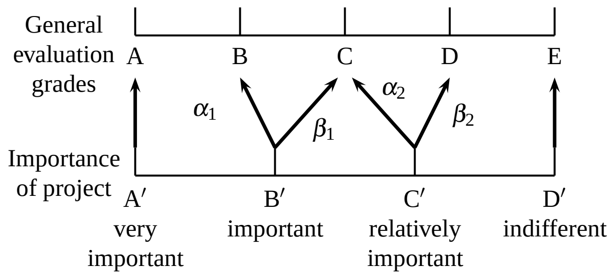

Fig. 4. Transformation Rules for Importance of project

In the above transformation formulae, $\alpha_1 + \beta_1 = 1$ and $\alpha_2 + \beta_2 = 1$, which mean the transformation is complete. The transformation matrix for "importance of project" is then given by

$$\overline{\mathrm{A}}(\text{importance}) = \begin{array}{c|ccccc} & \mathrm{A} & \mathrm{B} & \mathrm{C} & \mathrm{D} & \mathrm{E} \\ \mathrm{A}' & 1 & 0 & 0 & 0 & 0 \\ \mathrm{B}' & 0 & \alpha_1 & \beta_1 & 0 & 0 \\ \mathrm{C}' & 0 & 0 & \alpha_2 & \beta_2 & 0 \\ \mathrm{D}' & 0 & 0 & 0 & 0 & 1 \end{array} \quad (36)$$

Here, α_i $(i = 1, 2)$ and β_i $(i = 1, 2)$ are all assumed to be 0.5 in the case study.

d. Advance of critical techniques

The evaluation grades for "Advance of critical techniques" are defined according to the advance level of technology adopted in a product. The more advanced the technology adopted in a product, the better the product would be evaluated. Five evaluation grades are defined for "advance of critical technique", namely A' - very complex, B' - complex, C' - average, D' - poor, and E' - worst. They correspond to the 5 general evaluation grades with A' equivalent to A, B' to B, C' to C, D' to D and E' to E.

e. Ratio between quality and price

6 products of the same type in other 6 car manufacturers are selected and assessed on quality and price ratio. For example, when heavy trailer is assessed on quality and price ratio, 6 heavy trailers of the similar type (considering weight, oil consumption and so on) from other 6 car manufacturers are selected and evaluated on this factor. Then the quality and price ratio of the study heavy trailer is compared with the same type of products from the other 6 manufacturers as follows.

(i) If the study heavy trailer is much better than all the same type of other products, it would be assessed to the evaluation grade A'.

(ii) If it is better than most of other products and is only worse than one or two other products, it should be evaluated to B'.

(iii) If it is better than 3 other products but also worse than 3 other products, it should be evaluated to C'.

(iv) If it is worse than most of other products but better than one or two other products, it should be evaluated to D'.

(v) If it is worse than all other products, it should be assessed to E'.

f. Reliability of product

It is evaluated by the number of malfunction products in every 100 products in three months. It also needs to be transformed to the general evaluation grades.

g. Economy

It is evaluated by the amount of petrol consumed in every 100 miles (L/100miles).

h. Theoretical standard of project

It is assessed by the theoretical standard reflected from the R&D project. Five evaluation grades A', B', C', D' and E' are defined for this factor, in which A' represents the highest theoretical standard and E' the lowest theoretical standard. They all exactly correspond to the 5 general evaluation grades (A, B, C, D, and E).

i. Degree of innovation

Four evaluation grades are defined for the assessment of this factor as follows. They also need to be transformed to the general evaluation grades similar to Fig. 4, and equations (34), (35), (36):
A' - Leading internationally
B' - Filling a gap nationally
C' - Leading nationally
D' - Filling a gap for the company

j. Ratio of individual design

Three evaluation grades are defined for this factor. If a product is completely designed by the members of the company, it is assessed to the highest grade A'. If more than 60% of a product is designed by the members of the company, it is evaluated to B'. Otherwise, a product should be assessed to C'.

$$\mathrm{H}^{\text{individual design}} = \{\text{completely, more than 60\%, less than 60\%}\} \tag{37}$$

The three evaluation grades in this factor may also be transformed to the general five evaluation grades. From Fig. 5, it is clear that A', B' and C' each corresponds to A, C and E with a belief degree of 100%. It means that the utility of A' equals to the utility of A: $u(\mathrm{A}') = u(\mathrm{A})$; similarly $u(\mathrm{B}') = u(\mathrm{B})$ and $u(\mathrm{C}') = u(\mathrm{C})$. Different from factor b, there is no grade in factor j related to B and D in the general grade level.

Process control

k. Quality of project

It is evaluated on technical criteria. Three evaluation grades are defined. A project would be given the highest evaluation grade A′ if the project reaches the expected standard on all technical criteria. If accidents due to quality occur, a project would be assessed to the lowest grade C′. A project would be assessed to B′ if it does not reach the standard on some technical criteria. The transformation of these three evaluation grades to the general evaluation grades (A, B, C, D, and E) is similar to factor j.

l. Completion time for a project

It is evaluated by actual completion time compared with work schedule. If a R&D project is completed ahead of its schedule or on time, then as far as completion time is concerned, the project is evaluated to be the best (or A). In a similar way, the equivalence rule may also be stated as follows:

if a project is delayed by 2 months, the performance would be B;
if a project is delayed by 5 months, the performance would be C;
if a project is delayed by 7 months, the performance would be D;
if a project is delayed by 10 or more than 10 months, the performance would be E.

Table 1 may provide a straightforward relation between completion time and the general evaluation grades.

Table 1. Transformation Rules for Completion time for a project

Delayed months (h_i)	on time	2	5	7	More than 10
Evaluation grade (H_i)	A	B	C	D	E

By equivalence, the utility of h_i is equal to that of H_i. If the delayed time is not exactly the above 5 points associated with the 5 general evaluation grades, then it should be transformed to the 5 general grades similar to factor a and formula (30).

m. Investment of project

It is evaluated by the cost control compared with the company budget. A project would be assessed to the best grade A′ if there is cost saving or no over spending. If there is the overspending of 10%, it would be evaluated to grade B′. If the overspending is between 10% and 20%, it would be evaluated to grade C′. If the overspending is between 20% and 30%, it would be evaluated to grade D′. If the overspending is over 30% or the money for a project is defalcated, it would be evaluated to the worst grade E′. These five evaluation grades correspond to the general evaluation grades respectively.

Added value

n. Documents of rules and regulations established about project group

The following four evaluation grades are defined to assess this factor.

A′ - There are complete and excellent documents of rules and regulations, which can be used as a role model for other projects;
B′ - There are complete documents of rules and regulation, and the management of rules and regulation in a project is excellent;
C′ - There are documents of rules and regulation, and the management of rules and regulation in a project is indifferent;
D′ - There are no documents of rules and regulation.

These four evaluation grades could be transformed to the general grades according to the process defined in equations (34), (35) and (36).

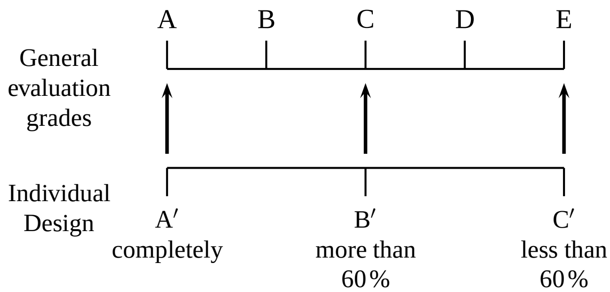

Fig. 5. Transformation Rules for Ration of Individual Design

o. Routine operational management documents

The following five evaluation grades are defined to assess this factor.

A′ - There are complete and excellent documents of routine management, which can be used as a role model for other projects;

- B′ - There are complete documents of routine management and the management of a project is good;
- C′ - There are documents of routine management and the management of a project is indifferent;
- D′ - There are no documents of routine management, but the management of a project is not too bad;
- E′ - There are no documents of routine management and the management and regulations are poor as well.

The evaluation grade of factor o could be transformed to the general evaluation grade defined in (4) respectively

p. Management documents about R&D process of products

This factor is assessed on the basis of the following five grades:

- A′ - There are complete and excellent documents of R&D process, which can be used as a role model for other projects;
- B′ - There are complete documents of R&D process and the management of R&D in a project is good;
- C′ - There are documents of R&D process and the management of R&D in a project is indifferent;
- D′ - There are no documents of R&D process, but the management of R&D in a project is not too bad;
- E′ - There are no documents of R&D process and the management of R&D is poor.

Each of the five evaluation grades for the assessment of factor p corresponds to the five general evaluation grades defined in (7) and (8) respectively.

q. Accumulation and continuity of technique

- A′ - New techniques applied in a R&D process have greatly contributed to future developments;
- B′ - There are limited or indifferent contributions to future developments by new techniques applied in a R&D project;
- C′ - New techniques applied in a R&D project have no contribution to future developments.

The transformation process in factor q is similar to factor j.

The overall assessment table is shown in Table 2 after the transformation of the original information (Tables 10 and 11) and the original belief degrees (Table 12) to the general evaluation grades. The transformation approach is as discussed above. The original information provided by experts is illustrated and transformed in the appendix.

5.3. *Acquiring the relative weight of factors through GAHP*

5.3.1. *Data collection*

The present project management and the evaluation system adopted by the company were investigated through many meetings including face-to-face discussions and interviews with its staff from each department. As a result of the investigation, the “standard table for checkout and assessment system of strategic R&D projects” is constructed. Apart from this onsite investigation, 200 questionnaires of the standard table were sent to the members of the R&D center and the relevant leaders of the company, from which 172 questionnaires were returned with 5 completely useless and 10 partly useless.

5.3.2. *Generate weights using the GAHP approach*

Considering the complexity of handling multiple criteria simultaneously, it is important to get a group of experts involved for assigning criteria weights, who should have different backgrounds and expertise and may represent conflicting interests. For such a group of experts, a question appears as to how to achieve group judgments from individual experts’ estimation of criteria weights. The GAHP approach is a method for generating the aggregated weight of a criterion from the judgments given by a group of experts.

At first, criteria which have great relevance to a project should be selected. Then, pairwise comparisons between every two criteria are provided by each expert to construct his/her judgmental matrix that is then used to generate a ranking vector by each expert. The judgmental matrix of each expert can be aggregated to form an overall judgmental matrix.

Table 2. The standard table for checkout and assessment system of strategic R&D project

General attributes	**Criteria in the second level**	**Factors in the lowest level (contents of assessment)**	**Type of project**			
			Light Trailer	**Heavy Trailer**	**MPV**	**SRV**
Quality of production E_1 ($\omega_1 = 0.40$)	Scale and importance of project E_{11} ($\omega_{11} = 0.3$)	workload e_{11}^1 ($\omega_{111} = 0.25$)	D(1.0)	D(1.0)	A(0.33) B(0.67)	A(1.0)
		origin of person e_{11}^2 ($\omega_{112} = 0.25$)	B(0.15) C(0.70) D(0.15)	B(0.15) C(0.70) D(0.15)	B(0.15) C(0.70) D(0.15)	B(0.15) C(0.70) D(0.15)
		importance of project e_{11}^3 ($\omega_{113} = 0.50$)	A(0.45) B(0.165) C(0.165) E(0.22)	A(0.83) B(0.085) C(0.085)	A(0.2) B(0.4) C(0.4)	A(0.70) B(0.15) C(0.15)
	Content of technique E_{12} ($\omega_{12} = 0.4$)	advance of critical techniques e_{12}^1 ($\omega_{121} = 0.1875$)	B(0.67) C(0.33)	A(0.67) B(0.33)	A(0.4) B(0.5) C(0.1)	A(0.6) B(0.3) C(0.1)
		ratio between quality and price e_{12}^2 ($\omega_{122} = 0.2500$)	A(0.56) B(0.22) C(0.22)	A(0.25) B(0.67) C(0.08)	A(0.3) B(0.7)	A(0.1) B(0.4) C(0.5)
		reliability of product e_{12}^3 ($\omega_{123} = 0.3750$)	H(1.0)	H(1.0)	H(1.0)	H(1.0)
		economy e_{12}^4 ($\omega_{124} = 0.1875$)	A(1.0)	B(1.0)	A(0.4) B(0.6)	B(0.7) C(0.3)
	Theoretical value and level of innovation E_{13} ($\omega_{13} = 0.3$)	theoretical standard of project e_{13}^1 ($\omega_{131} = 0.25$)	A(0.11) B(0.45) C(0.33) D(0.11)	A(0.42) B(0.50) D(0.08)	A(0.5) B(0.5)	A(0.8) B(0.2)
		degree of innovation e_{13}^2 ($\omega_{132} = 0.42$)	B(0.05) C(0.275) D(0.225) E(0.45)	B(0.125) C(0.415) D(0.29) E(0.17)	A(0.1) B(0.1) C(0.35) D(0.25) E(0.2)	B(0.1) C(0.2) D(0.1) E(0.6)
		ratio of individual design e_{13}^3 ($\omega_{133} = 0.33$)	B(1.0)	B(1.0)	B(1.0)	B(1.0)
Process control E_2 ($\omega_2 = 0.35$)	Quality of project E_{21} ($\omega_{21} = 0.51$)		A(0.44) B(0.56)	A(0.83) B(0.17)	A(0.5) B(0.4) C(0.1)	A(0.5) B(0.5)
	Completion time for a project E_{22} ($\omega_{22} = 0.20$)		B(1.0)	A(1.0)	A(1.0)	A(1.0)
	Investment E_{23} ($\omega_{23} = 0.29$)		A(1.0)	A(1.0)	A(1.0)	A(1.0)
Added value by project E_3 ($\omega_3 = 0.25$)	Project team E_{31} ($\omega_{31} = 0.48$)	documents of rules and regulations established about project group e_{31}^1 ($\omega_{311} = 0.25$)	A(0.22) B(0.11) C(0.335) D(0.225) E(0.11)	A(0.33) B(0.21) C(0.335) D(0.125)	A(0.2) B(0.2) C(0.4) D(0.2)	A(0.2) B(0.2) C(0.4) D(0.2)
		routine operational management documents e_{31}^2 ($\omega_{312} = 0.33$)	A(0.11) B(0.33) C(0.56)	A(0.25) B(0.58) C(0.17)	A(0.3) B(0.5) C(0.2)	A(0.1) B(0.5) C(0.4)
		management documents about R&D process of products e_{31}^3 ($\omega_{313} = 0.42$)	A(0.22) B(0.11) C(0.56) D(0.11)	A(0.33) B(0.58) C(0.09)	A(0.6) B(0.3) C(0.1)	A(0.4) B(0.4) C(0.2)
	Continuity of technique E_{32} ($\omega_{32} = 0.52$)	accumulation and continuity of technique	A(0.67) C(0.22) E(0.11)	A(0.92) E(0.08)	A(0.7) C(0.2) E(0.1)	A(0.7) C(0.3)

There are some approaches for the aggregation of matrices, for instance, the ideal synthesis matrix in a group context, additive convex set and Hadamard convex set of judgmental matrix.[6] These approaches can minimize the impact of inconsistence among individual expert judgments. In this paper, Hadamard convex set of judgmental matrix is implemented to construct an overall pairwise matrix. Then, the analytic hierarchy process (AHP) approach is applied to calculate the weight of each criterion. This whole process is referred to as a generalized AHP approach, or GAHP. The detailed calculation process and Hadamard convex set of a judgmental matrix are described in the appendix.

Based on the 172 questionnaires, the GAHP is used to calculate the weights of all attributes used in the checkout and assessment system as follows. Weights of attributes in the first level are given by:

$$\omega_1 = 0.3957, \quad \omega_2 = 0.3617, \quad \omega_3 = 0.2426;$$

Weights of attributes in the second level are given by:

$$\omega_{11} = 0.2886, \quad \omega_{12} = 0.3845, \quad \omega_{13} = 0.3269;$$
$$\omega_{21} = 0.4919, \quad \omega_{22} = 0.2175, \quad \omega_{23} = 0.2906;$$
$$\omega_{31} = 0.4794, \quad \omega_{32} = 0.5206;$$

Weights of attributes in the third level are given by:

$$\omega_{111} = 0.2247,\ \omega_{112} = 0.2438,\ \omega_{113} = 0.5315;$$
$$\omega_{121} = 0.1319,\ \omega_{122} = 0.2692,\ \omega_{123} = 0.3781,\ \omega_{124} = 0.2208;$$
$$\omega_{131} = 0.2446,\ \omega_{132} = 0.4050,\ \omega_{133} = 0.3504;$$
$$\omega_{311} = 0.2842,\ \omega_{312} = 0.3430,\ \omega_{313} = 0.3728;$$

Since the company is operating in a dynamic business environment, the above weights could be revised in the future. The values of weights are rounded, as shown in Table 2.

Table 3. Subjective assessment about theoretical value and innovation of Heavy Trailer

Weights	Attributes	Evaluation grades				
		E	D	C	B	A
$\omega_{131} = 0.25$	Theoretical standard	0	0.08	0	0.50	0.42
$\omega_{132} = 0.42$	Degree of Innovation	0.17	0.29	0.415	0.125	0
$\omega_{133} = 0.33$	Ratio of individual design	0	0	0	1.0	0

5.4. ER modeling framework for R&D project assessment

When the ER approach is applied to generate the aggregated assessments for R&D projects, e_{11}^1 to e_{11}^3, e_{12}^1 to e_{12}^4, e_{13}^1 to e_{13}^3 and e_{31}^1 to e_{31}^3 are aggregated first, resulting in the assessments of the first level attributes. Then E_{11} to E_{13}, E_{21} to E_{23} and E_{31} to E_{32} are aggregated to generate the assessments for these second level attributes. Finally, the three third level attributes: quality of production, process control and added value are aggregated to generate the overall assessment for each project. This forms a systematic aggregating process for the generation of overall assessment for a R&D project.

For example, the criterion "theoretical value and level of innovation" is assessed through the three lower level attributes: theoretical standard of project, degree of innovation, ratio of individual design. The assessments of the heavy trailer project produced by the company on these three attributes are shown in Table 3.

We could see that these three assessments are all complete. From Table 3 we have:

$$\beta_{11} = 0, \beta_{21} = 0.08, \beta_{31} = 0, \beta_{41} = 0.50, \beta_{51} = 0.42,$$
$$\beta_{12} = 0.17, \beta_{22} = 0.29, \beta_{32} = 0.415, \beta_{42} = 0.125, \beta_{52} = 0,$$
$$\beta_{13} = 0, \beta_{23} = 0, \beta_{33} = 0, \beta_{43} = 1.0, \beta_{53} = 0.$$

From the belief degrees and the weights calculated using GAHP, we could have the following basic probability masses:

$m_{11} = 0, \quad m_{21} = 0.02, \quad m_{31} = 0,$
$m_{41} = 0.125, \quad m_{51} = 0.105,$

$$\overline{m}_{H,1} = 1 - \omega_{131} = 0.75,$$
$$\widetilde{m}_{H,1} = \omega_{131}\left(1 - \sum_{n=1}^{5} \beta_{n1}\right) = 0$$

$m_{12} = 0.0714, \quad m_{22} = 0.1218, \quad m_{32} = 0.1743,$
$m_{42} = 0.0525, \quad m_{52} = 0.$

$$\overline{m}_{H,2} = 1 - \omega_{132} = 0.58,$$
$$\widetilde{m}_{H,2} = \omega_{132}\left(1 - \sum_{n=1}^{5} \beta_{n2}\right) = 0$$

$m_{13} = 0, m_{23} = 0, m_{33} = 0, m_{43} = 0.33, m_{53} = 0.$

$$\overline{m}_{H,3} = 1 - \omega_{133} = 0.67,$$
$$\widetilde{m}_{H,3} = \omega_{133}\left(1 - \sum_{n=1}^{5} \beta_{n3}\right) = 0$$

Firstly, the assessments on "theoretical standard" and "level of innovation" are aggregated to generate an intermediate assessment on $E_{I(2)}$, as shown by equations (18), (19), (20), and (21),

$$k_{I(2)} = \left[1 - \sum_{s=1}^{5} \sum_{\substack{j=1 \\ j \neq s}}^{5} m_{s,1} m_{j,2}\right]^{-1} = 1.1062$$

$$\begin{aligned}
m_{1,I(2)} &= k_{I(2)}(m_{11}m_{12} + m_{11}m_{H,2} + m_{12}m_{H,1}) \\
&= 0.05924 \\
m_{2,I(2)} &= k_{I(2)}(m_{21}m_{22} + m_{21}m_{H,2} + m_{22}m_{H,1}) \\
&= 0.11658 \\
m_{3,I(2)} &= k_{I(2)}(m_{31}m_{32} + m_{31}m_{H,2} + m_{32}m_{H,1}) \\
&= 0.14461 \\
m_{4,I(2)} &= k_{I(2)}(m_{41}m_{42} + m_{41}m_{H,2} + m_{42}m_{H,1}) \\
&= 0.13102 \\
m_{5,I(2)} &= k_{I(2)}(m_{51}m_{52} + m_{51}m_{H,2} + m_{52}m_{H,1}) \\
&= 0.067367 \\
\overline{m}_{H,I(2)} &= k_{I(2)}\overline{m}_{H,1}\overline{m}_{H,2} = 0.4812 \\
\widetilde{m}_{H,I(2)} &= k_{I(2)}(\widetilde{m}_{H,1}\widetilde{m}_{H,2} + \overline{m}_{H,1}\widetilde{m}_{H,2} + \widetilde{m}_{H,1}\overline{m}_{H,2}) = 0
\end{aligned}$$

Then, generate the assessment on "theoretical and innovation level" ($E_{I(3)}$) by aggregating $E_{I(2)}$ and the assessment on individual design:

$$k_{I(3)} = \left[1 - \sum_{s=1}^{5} \sum_{\substack{j=1 \\ j \neq s}}^{5} m_{s,I(2)} m_{j,3}\right]^{-1} = 1.14675$$

$$\begin{aligned}
m_{1,I(3)} &= k_{I(3)}(m_{1,I(2)}m_{13} + m_{1,I(2)}m_{H,3} + m_{13}m_{H,I(2)}) \\
&= 0.04551 \\
m_{2,I(3)} &= k_{I(3)}(m_{2,I(2)}m_{23} + m_{2,I(2)}m_{H,3} + m_{23}m_{H,I(2)}) \\
&= 0.08957 \\
m_{3,I(3)} &= k_{I(3)}(m_{3,I(2)}m_{33} + m_{3,I(2)}m_{H,3} + m_{33}m_{H,I(2)}) \\
&= 0.11111 \\
m_{4,I(3)} &= k_{I(3)}(m_{4,I(2)}m_{43} + m_{4,I(2)}m_{H,3} + m_{43}m_{H,I(2)}) \\
&= 0.33234 \\
m_{5,I(3)} &= k_{I(3)}(m_{5,I(2)}m_{53} + m_{5,I(2)}m_{H,3} + m_{53}m_{H,I(2)}) \\
&= 0.05176 \\
\overline{m}_{H,I(3)} &= k_{I(3)}\overline{m}_{H,I(2)}\overline{m}_{H,3} = 0.3697 \\
\widetilde{m}_{H,I(3)} &= k_{I(3)}(\widetilde{m}_{H,I(2)}\widetilde{m}_{H,3} + \overline{m}_{H,I(2)}\widetilde{m}_{H,3} \\
&\qquad + \widetilde{m}_{H,I(2)}\overline{m}_{H,3}) = 0
\end{aligned}$$

Then, the belief degrees of "theoretical value and level of innovation" could be calculated from (22) and (23),

$$\beta_1 = \frac{m_{1,I(3)}}{1 - \overline{m}_{H,I(3)}} = 0.0722, \quad \beta_2 = \frac{m_{2,I(3)}}{1 - \overline{m}_{H,I(3)}} = 0.1421,$$
$$\beta_3 = \frac{m_{3,I(3)}}{1 - \overline{m}_{H,I(3)}} = 0.1763, \quad \beta_4 = \frac{m_{4,I(3)}}{1 - \overline{m}_{H,I(3)}} = 0.5273,$$
$$\beta_5 = \frac{m_{5,I(3)}}{1 - \overline{m}_{H,I(3)}} = 0.0821, \quad \beta_H = \frac{\widetilde{m}_{H,I(3)}}{1 - \overline{m}_{H,I(3)}} = 0,$$

The assessment of the heavy trailer project on the attribute can be represented by the following statement:

$$S(\text{theoretical value and level of innovation}) = \{(\text{Worst}, 0.0722), (\text{Poor}, 0.1421), (\text{Average}, 0.1763), (\text{Good}, 0.5273), (\text{Best}, 0.0821)\}$$

The overall assessment of the heavy trailer can be aggregated in the same fashion, given by:

$\beta_1 = 0.0143, \quad \beta_2 = 0.0400, \quad \beta_3 = 0.0538$
$\beta_4 = 0.2185, \quad \beta_5 = 0.6361, \quad \beta_H = 0.0374.$

So the distributed assessment of the heavy trailer project is represented as follows:

$$S(\text{Heavy Trailer}) = \{(\text{Worst}, 0.0143), (\text{Poor}, 0.0400), (\text{Average}, 0.0538), (\text{Good}, 0.2185), (\text{Best}, 0.6361), (\text{H}, 0.0374)\}$$

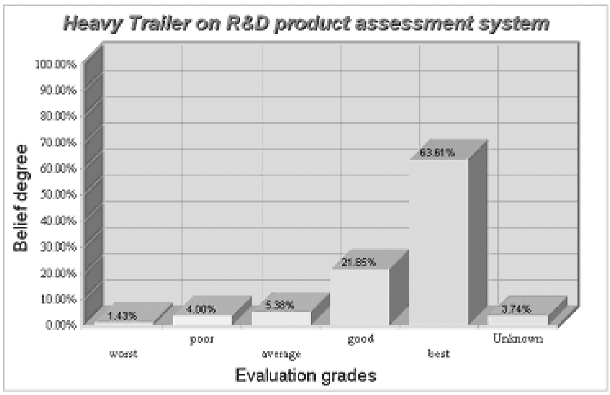

Fig. 6.

The incompleteness in the assessment of the heavy trailer is measured by a belief degree of 0.0374 due to the complete ignorance in the attribute “reliability of product”. It is the same case to the assessments of the light trailer, MPV and SRV projects. Using the IDS software, the assessment of the heavy trailer project on each grade can be visually presented as shown in Fig. 6.

Similarly, the assessments of the Light Trailer, MPV and SRV projects can also be generated as follows:

$$S(\text{Light Trailer}) = \{(\text{Worst}, 0.0468), (\text{Poor}, 0.0231), (\text{Average}, 0.1529), (\text{Good}, 0.3346), (\text{Best}, 0.4021), (\text{H}, 0.0406)\}$$

$$S(\text{MPV}) = \{(\text{Worst}, 0.0173), (\text{Poor}, 0.0159), (\text{Average}, 0.1126), (\text{Good}, 0.2951), (\text{Best}, 0.5220), (\text{H}, 0.0370)\}$$

$$S(\text{SRV}) = \{(\text{Worst}, 0.0266), (\text{Poor}, 0.0100), (\text{Average}, 0.1297), (\text{Good}, 0.2375), (\text{Best}, 0.5570), (\text{H}, 0.0391)\}$$

The above distributed assessments provide a panoramic view of the overall performances of each of these different projects individually. However, it is not easy to rank the four projects just based on the distributed assessments. The utility intervals can be calculated for the comparison. Suppose the utility of the evaluation grades are equidistantly distributed according to equation (14). Then the maximum utility, minimum utility and average utility of the heavy trailer project can be calculated according to equations (25), (26), (27) and (28) as follows:

$$u(S(y(\text{Heavy Trailer}))) = \sum_{n=1}^{5} \beta_n \times u(\text{H}_n) = 0.8369$$

$$u_{\max}(\text{Heavy Trailer}) = u(S(y(\text{Heavy Trailer}))) + \beta_{\text{H}}(\text{Heavy Trailer})u(\text{H}_5) = 0.8742$$

$$u_{\min}(\text{Heavy Trailer}) = u(S(y(\text{Heavy Trailer}))) + \beta_{\text{H}}(\text{Heavy Trailer})u(\text{H}_1) = 0.8369$$

$$u_{\text{ave}}(\text{Heavy Trailer}) = \frac{u_{\max}(\text{Heavy Trailer}) + u_{\min}(\text{Heavy Trailer})}{2} = \sum_{n=1}^{5} \beta_n \times u(\text{H}_n) + \beta_{\text{H}}(\text{Heavy Trailer})\frac{u(\text{H}_5) + u(\text{H}_1)}{2} = 0.8556$$

Similarly, the utility intervals for the other three projects could also be calculated as shown in Table 4.

Table 4. Utility interval of the four R&D projects

	Light Trailer	Heavy Trailer	MPV	SRV
Maximum Utility	0.7758	0.8742	0.8406	0.8416
Minimum Utility	0.7353	0.8369	0.8036	0.8025
Average Utility	0.7556	0.8556	0.8221	0.8221
Rank	4	1	2	2

From the utility interval, we could get the ranking order of the four projects using formula (29):

$$P(\text{Heavy Trailer} \succ \text{MPV}) = \min\left\{\max\left(\frac{0.8742-0.8036}{[0.8742-0.8369]+[0.8406-0.8036]}, 0\right), 1\right\} = 95.02\%$$

$$P(\text{MPV} \succ \text{SRV}) = \min\left\{\max\left(\frac{0.8406-0.8025}{[0.8406-0.8036]+[0.8416-0.8025]}, 0\right), 1\right\} = 50.07\%$$

From Table 4, it is clear that the maximum utility of the light trailer project is less than the minimum utility of all the other three products, so the light trailer project is ranked the last. It is also obvious that the heavy trailer project is superior to the MPV project to the extent of 95.02%, whereas the MPV project is superior to the SRV project to the extent of 50.07%. Also, the average utility of MPV almost equals that of the SRV. So, these two R&D projects are almost equivalent. The four projects could be ranked as follows with certain level of confidence:

$$\text{Heavy Trailer} \succ \text{MPV} \sim \text{SRV} \succ \text{Light Trailer}$$

Here, the ranking order of the four projects generated by formula (29) is the same as that calculated by average utility, though this is not always the case. Figure 7 and Table 5 shows the average utilities of these four R&D projects on the overall performance and three second level attributes. From the figure, it is clear that the average utility of light trailer is obvious inferior to the other three projects, and heavy trailer is superior evidently.

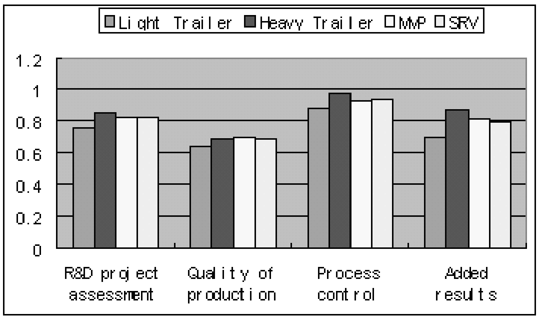

Fig. 7. Average utilities of the four R&D projects

Although the combined weights calculated using the GAHP methods represent the aggregated opinion of all experts involved in the investigation, they do not necessarily reflect the views of individual experts or departments. So the sensitivity analysis of weights should be conducted based on the above combined weights. In the investigation process in the car manufacturer, 200 questionnaires were handed out to different departments as mentioned in Section 5. The three different departments involved are the marketing department, manufacturing department, and human resource department of the car manufacturer. GAHP is used for the data collected from two of the three departments and two different pieces of weights are generated afterwards. The following are the rounded values of the weights calculated using GAHP. Weights calculated using the information collected from the marketing department are as follows.

Table 5. The average utilities of the four R&D projects on major attributes

R&D projects	General assessment	Quality of production	Process control	Added results
Light Trailer	0.7556	0.6438	0.879607	0.699717
Heavy Trailer	0.8556	0.687833	0.982007	0.870727
MPV	0.8221	0.696884	0.928944	0.816872
SRV	0.8221	0.689781	0.940787	0.792692

The weights of the attributes in the first level are given by:

$$\omega_1 = 0.40, \quad \omega_2 = 0.44, \quad \omega_3 = 0.16;$$

The weights of the attributes in the second level are given by:

$$\omega_{11} = 0.30, \quad \omega_{12} = 0.60, \quad \omega_{13} = 0.10;$$
$$\omega_{21} = 0.11, \quad \omega_{22} = 0.53, \quad \omega_{23} = 0.36;$$
$$\omega_{31} = 0.45, \quad \omega_{32} = 0.55;$$

The weights of the attributes in the third level are given by:

$$\omega_{111} = 0.50, \quad \omega_{112} = 0.25, \quad \omega_{113} = 0.25;$$
$$\omega_{121} = 0.13, \quad \omega_{122} = 0.46, \quad \omega_{123} = 0.11, \quad \omega_{124} = 0.30;$$
$$\omega_{131} = 0.25, \quad \omega_{132} = 0.42, \quad \omega_{133} = 0.33;$$
$$\omega_{311} = 0.25, \quad \omega_{312} = 0.35, \quad \omega_{313} = 0.40;$$

The weights calculated using the information collected from the manufacturing department are as follows. The weights of the attributes in the first level are given by:

$$\omega_1 = 0.47, \quad \omega_2 = 0.34, \quad \omega_3 = 0.19$$

Table 6. The utility of each R&D products under the weights calculated from marketing department

	Light Trailer	Heavy Trailer	MPV	SRV
Maximum utility	0.820307	0.882935	0.908186	0.882412
Minimum utility	0.806130	0.868959	0.895613	0.867977
Average utility	0.813218	0.875947	0.901899	0.875194
Rank	4	2	1	3

The weights of the attributes in the second level are given by:

$$\omega_{11} = 0.41, \quad \omega_{12} = 0.35, \quad \omega_{13} = 0.24;$$
$$\omega_{21} = 0.60, \quad \omega_{22} = 0.28, \quad \omega_{23} = 0.12;$$
$$\omega_{31} = 0.49, \quad \omega_{32} = 0.51;$$

The weights of the attributes in the third level are given by:

$$\omega_{111} = 0.52, \quad \omega_{112} = 0.23, \quad \omega_{113} = 0.25;$$
$$\omega_{121} = 0.22, \quad \omega_{122} = 0.30, \quad \omega_{123} = 0.17,$$
$$\omega_{124} = 0.31;$$
$$\omega_{131} = 0.28, \quad \omega_{132} = 0.40, \quad \omega_{133} = 0.32;$$
$$\omega_{311} = 0.34, \quad \omega_{312} = 0.39, \quad \omega_{313} = 0.27;$$

Light trailer is ranked the worst among the four R&D projects no matter whether the ranking is based on the weights calculated by the marketing department, the manufacturing department or all the departments. Heavy trailer is ranked the best one among these four R&D projects using the weights calculated from the information provided by all the departments, while MPV is ranked the best one under the information provided by the marketing department and SRV is ranked the best from the information provided by the manufacturing department. The ranking results are inconsistent due to the different opinions to the importance of the attributes from people in the different departments. So it is significant to construct several sets of weights from different departments to support the decision making process.

Table 7. The utility of each R&D products under the weights calculated from manufacturing department

	Light Trailer	Heavy Trailer	MPV	SRV
Maximum utility	0.762167	0.781756	0.825977	0.846581
Minimum utility	0.746213	0.765699	0.811814	0.830847
Average utility	0.754190	0.773728	0.818895	0.838714
Rank	4	3	2	1

6. Concluding remarks

In this paper, we applied the evidential reasoning approach to the assessment of strategic R&D projects for a car manufacturer. A strategic R&D project in the car manufacturing industry is normally concerned with lots of money and human resources and lasts a long period of time. From the case study, a reliable and rational hierarchy of attributes for the assessment of R&D projects was constructed. The GAHP method is introduced to calculate the weights of the attributes in the evaluation hierarchy by synthesizing 100 experts opinions with different backgrounds and knowledge. For acquiring the original information represented by distributed evaluation grades on qualitative attributes (Table 11) and numerical values on quantitative attributes (Table 10) of each R&D project from experts in the car manufacturer, one table for assessing quantitative attributes was provided to one expert in each R&D project, and tables for assessing qualitative attributes were provided to a group of experts in each R&D project. The distributed assessment of each R&D project on each attribute (Table 12) is calculated by means of rules as investigated in 5.2.2 and frequency as mentioned in appendix A.3, which provides a vigorous, flexible yet pragmatic way for transforming surveyed data decision knowledge for decision modeling and analysis. The evidential reasoning approach is suitable to assessing the R&D projects for car manufacturers, which is characterized by the inherent uncertainty of human judgments that exist in the R&D project assessment process. Although the GAHP method can be used to combine opinions from a group of experts in terms of the calculation of weights, such aggregated weights may not represent the different views of individual experts or departments. As such, the sensitivity analysis on weights was conducted in the paper to examine the impact of different weights on the assessment and ranking of projects. In this paper, only preliminary sensitivity analysis was conducted. In future research, more comprehensive analyses on interval or fuzzy weights will need to be conducted to support more informative decision making.

Acknowledgements

This work is supported by the key program of National Natural Science Foundation of China (NSFC) under the Grant No 70631003 and the National Hi-Tech Research and Development Program of China (863 Program) under the Grant No 2006AA04Z134.

Appendix A.

In the appendix, the convex set of Hadamard multiplication and the calculation process of GAHP are outlined, and the original information collected is also presented.

A.1. Convex set of Hadamard multiplication

(see [6])

Definition 1. The Hadamard multiplication of matrix $A = (a_{ij})_{n\times n}$ and matrix $\mathrm{B} = (b_{ij})_{n\times n}$ is defined as follows, represented by $C = (c_{ij})_{n\times n}$:

$$c_{ij} = a_{ij} \cdot b_{ij} \tag{A.1}$$

which is denoted by $C = A \cdot B$.

Definition 2. Suppose $\alpha \in \mathbb{R}$. The exponential algorithm of matrix $A = (a_{ij})_{n\times n}$ is defined as:

$$A^{\alpha} = \left(a_{ij}^{\alpha}\right)_{n\times n} \tag{A.2}$$

Definition 3. If $a_{ij} = a_{ji}^{-1}$, $i \neq j$, $i, j = 1, 2, \ldots, n$, then matrix $A = (a_{ij})_{n\times n}$ is called reflexive matrix.

Definition 4. Suppose $A_1, A_2, \ldots, A_m$ are m judgmental matrices for the same problem. If $\lambda_1, \lambda_2, \ldots, \lambda_m$ exist, λ_i $(i = 1, 2, \ldots, m)$ satisfy $\lambda_i \in [0,1]$ and $\sum_{i=1}^{m} \lambda_i = 1$, and $\overline{A} = A_1^{\lambda_1} \cdot A_2^{\lambda_2} \cdot \cdots \cdot A_m^{\lambda_m}$, then $\overline{A} = (\overline{a}_{ij})_{n\times n}$ is called a convex set of Hadamard multiplication from $A_1, A_2, \ldots, A_m$, where

$$\overline{a}_{ij} = \left(a_{ij}^{(1)}\right)^{\lambda_1} \cdot \left(a_{ij}^{(2)}\right)^{\lambda_2} \cdot \cdots \cdot \left(a_{ij}^{(m)}\right)^{\lambda_m}, \quad (i, j = 1, 2, \ldots, m). \tag{A.3}$$

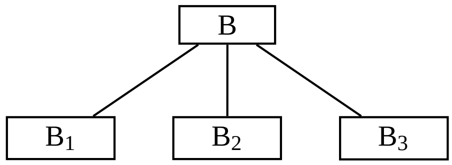

Fig. 8. Hierarchical structure of the attributes

Suppose A^{*} is the completely consistent matrix. From $A_L = A^{*}\varepsilon_L$ $(L = 1, 2, \ldots, m)$, we can generate ε_L as follow:

$$\varepsilon_L = \begin{pmatrix} \varepsilon_{11}^{(L)} & \cdots & \varepsilon_{1n}^{(L)} \\ \cdots & \cdots & \cdots \\ \varepsilon_{n1}^{(L)} & \cdots & \varepsilon_{nn}^{(L)} \end{pmatrix} \quad (L = 1, 2, \ldots, m),$$

where the elements are non-negative, and $\varepsilon_{ij} = \varepsilon_{ji}^{-1}$ $(i, j = 1, 2, \ldots, n)$.

Suppose

$$\varepsilon = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \tag{A.4}$$

If $\varepsilon_L = \varepsilon$, then A_L is supposed to be a consistent matrix of a matrix. We denote $\overline{A} = A^{*} \cdot \overline{\varepsilon}$, with $\overline{\varepsilon} = (\overline{\varepsilon}_{ij})_{n\times n}$, where

$$\overline{\varepsilon}_{ij} = \left(\varepsilon_{ij}^{(1)}\right)^{\lambda_1} \cdot \left(\varepsilon_{ij}^{(2)}\right)^{\lambda_2} \cdot \cdots \cdot \left(\varepsilon_{ij}^{(m)}\right)^{\lambda_m}, \quad (i, j = 1, 2, \ldots, n) \tag{A.5}$$

Theorem 1. *Suppose $A_1, A_2, \ldots, A_m$ are m judgmental matrices for the same problem. Parameter $R > 0$ is small enough that leads to the equation*

$$\frac{1}{n(n-1)} \sum_{1 \leqslant i < j \leqslant n} \left| \log \varepsilon_{ij}^{(L)} \right| \leqslant R \quad (L = 1, 2, \ldots, m)$$

and $\overline{A}$ is a convex set of Hadamard multiplication from $A_1, A_2, \ldots, A_m$. Then we will have the conclusion:

$$\frac{1}{n(n-1)} \sum_{1 \leqslant i < j \leqslant n} \left| \log \overline{\varepsilon}_{ij} \right| \leqslant R,$$

where $\frac{1}{n(n-1)} \sum_{1 \leqslant i < j \leqslant n} \left| \log \overline{\varepsilon}_{ij} \right|$ *is the measurement of the consistence of $\overline{A}$.*

A.2. Step of GAHP

In the following, a simple example is discussed to illustrate the calculation process of the GAHP approach. The details about the AHP approach can be found in Ref. [8,9]. A two level attribute hierarchy is presented in Fig. 8, in which three basic attributes are associated with one top level attribute. For illustration purpose, five experts out of the 172 experts contacted in our survey are selected for constructing the pairwise comparison matrices of the three basic attributes. They are supposed to be equally important, or $\lambda_i = 0.2$ ($\lambda_i = 1, \ldots, 5$).

In this paper, we use a three-level judgmental scale.[26] If B_i is as important as B_j, the judgmental scale is 1; if B_i is more important than B_j, then it is 2; if B_i is less important than B_j, then it is 0. This $[0,1,2]$ three-level scale is simpler to use by practitioners than the 1-9 nine-level scale proposed by Saaty.[8,9] Based on Fig. 8, and the three-level scale, each expert is asked to give his/her own comparison matrix, leading to the following comparison matrices accompanied by consistency inspection. The comparison matrices provided by the five experts are shown in Table 8.

Table 8. Comparison matrices by five experts

Expert 1:

B	B_1	B_2	B_3	*W*	B	B_1	B_2	B_3
B_1	1	1	2		B_1	1	1	5
B_2	1	1	2		B_2	1	1	5
B_3	0	0	5		B_3	1/5	1/5	1

$\lambda_{\max} = 3,\ CR = 0 < 0.1$

Expert 2:

B	B_1	B_2	B_3	*W*	B	B_1	B_2	B_3
B_1	1	2	2		B_1	1	3	5
B_2	0	1	2		B_2	1/3	1	3
B_3	0	0	1		B_3	1/5	1/3	1

$\lambda_{\max} = 3.039,\ CR = 0.0332 < 0.1$

Expert 3:

B	B_1	B_2	B_3	*W*	B	B_1	B_2	B_3
B_1	1	0	0		B_1	1	1/3	1/5
B_2	2	1	0		B_2	3	1	1/3
B_3	2	2	1		B_3	5	3	1

$\lambda_{\max} = 3.039,\ CR = 0.0332 < 0.1$

Expert 4:

B	B_1	B_2	B_3	*W*	B	B_1	B_2	B_3
B_1	1	1	2		B_1	1	1	5
B_2	1	1	2		B_2	1	1	5
B_3	0	0	1		B_3	1/5	1/5	1

$\lambda_{\max} = 3,\ CR = 0 < 0.1$

Expert 5:

B	B_1	B_2	B_3	*W*	B	B_1	B_2	B_3
B_1	1	2	2		B_1	1	3	5
B_2	0	1	2		B_2	1/3	1	3
B_3	0	0	1		B_3	1/5	1/3	1

$\lambda_{\max} = 3.039,\ CR = 0.3332 < 0.1$

On the basis of the above comparison matrices from the five experts, we can generate the aggregated judgmental matrices (Table 6) via the convex set of Hadamard multiplication as mentioned in appendix A.1. The $[0,1,2]$ scale should be transformed to the 1-9 scale before the aggregation process. The right sides of Table 5 are based on the 1-9 scale after transformation. The consistency ratio of the aggregated matrices could also be calculated.

Let $a_{ij}^{(m)}$ be the comparison coefficient of B_i over B_j provided by expert m under the $[0,1,2]$ scale, and it is transformed to ${a'_{ij}}^{(m)}$ under the 1-9 scale. Let $\overline{a}_{ij}$ be the aggregated comparison coefficient of B_i over B_j. From Definition 4, we have

$$\overline{a}_{ij} = \begin{cases} \left({a'_{ij}}^{(1)}\right)^{0.2} \cdot & \left({a'_{ij}}^{(2)}\right)^{0.2} \cdot \cdots \cdot \left({a'_{ij}}^{(m)}\right)^{0.2} \\ & (i \neq j;\ i,\ j = 1, 2, 3) \\ 1 & i = j \end{cases} \tag{A.6}$$

So, $\bar{a}_{12} = 1^{0.2} \cdot 3^{0.2} \cdot (1/3)^{0.2} \cdot 1^{0.2} \cdot 3^{0.2} = 1.2457,$

$$\bar{a}_{21} = \frac{1}{\bar{a}_{12}} = 0.8027,$$
$$\bar{a}_{13} = 5^{0.2} \cdot 5^{0.2} \cdot (1/5)^{0.2} \cdot 5^{0.2} \cdot 5^{0.2} = 2.6265,$$
$$\bar{a}_{31} = \frac{1}{\bar{a}_{13}} = 0.3807,$$
$$\bar{a}_{23} = 5^{0.2} \cdot 3^{0.2} \cdot (1/3)^{0.2} \cdot 5^{0.2} \cdot 3^{0.2} = 2.3714,$$
$$\bar{a}_{32} = \frac{1}{\bar{a}_{23}} = 0.4217.$$

Table 9. Aggregated judgmental matrices

B	B_1	B_2	B_3	W
B_1	1	1.2457	2.6265	0.4543
B_2	0.8027	1	2.3741	0.37940
B_3	0.3807	0.4217	1	0.1663

$\lambda_{\max} = 3.001,\ CR = 0.0016 < 0.1$

The total consistency ratio (CR) is calculated as $CR = 0.0016 < 0.1$. So the weights of the general attributes are acquired as follows:

$$\omega_B = (\omega_{B1}, \omega_{B2}, \omega_{B3}) = (0.45, 0.38, 0.17)$$

In essence, GAHP is a process of using AHP to calculate the judgmental matrices and weights for each expert individually, and then using the convex set of Hadamard multiplication to aggregate these inconsistency data for generating overall matrices and weights. The results represent the judgment of the whole group.

A.3. The original information collected

In the process of collecting the assessment information about the R&D projects, 17 attributes in the assessment hierarchy are split into two parts: 7 quantitative and 10 qualitative attributes. In assessing these four R&D projects on the 7 quantitative attributes, each project is assessed by an expert who participated in the R&D process and is quite familiar with it. From Table 10, it is clear that there is no assessment for "reliability of product (e_{12}^3)" on each R&D project.

When assessing the four projects on the 10 qualitative attributes, it is not enough for only one expert to assess each R&D project due to the subjective nature of the assessments. A group of experts who joined in the process of R&D in a project were invited to assess the project on these 10 attributes. 9 experts were invited for assessing the light trailer project, 12 experts for heavy trailer, 10 experts for MPV, and 10 experts for SRV. Table 11 shows the original assessment data given by the groups of experts on the 10 qualitative attributes.

Let $P(a_l)$ denotes the number of experts who take part in the assessment of R&D project a_l, and $P_{H_{n,i}}(a_l)$ denotes the number of experts who assess a_l to the n^{th} evaluation degree on the i^{th} qualitative attribute. Take e_{11}^3 (importance of project) for example. The belief degree of light trailer assessed to A$'$ on e_{11}^3 could then be calculated as follows:

$$\beta_{\mathrm{A}',1}(\text{light trailer}) = \frac{P_{H_{4,1}}(\text{light trailer})}{P(\text{light trailer})} = \frac{4}{9} = 0.45$$

Similarly,

$$\beta_{\mathrm{B}',1}(\text{light trailer}) = \frac{P_{H_{3,1}}(\text{light trailer})}{P(\text{light trailer})} = \frac{3}{9} = 0.33$$
$$\beta_{\mathrm{C}',1}(\text{light trailer}) = \frac{P_{H_{2,1}}(\text{light trailer})}{P(\text{light trailer})} = \frac{0}{9} = 0$$
$$\beta_{\mathrm{D}',1}(\text{light trailer}) = \frac{P_{H_{1,1}}(\text{light trailer})}{P(\text{light trailer})} = \frac{2}{9} = 0.22$$

The assessments of the four projects on the other 9 qualitative attributes can be calculated similarly. Table 12 shows the belief degree after the calculation process.

Table 10. Original assessment in quantitative attributes

Quantitative attributes		(1) e^1_{11} (Months)	(2) e^2_{11}	(3) e^3_{12}	(4) e^4_{12} (L/100km)	(5) e^3_{13}	(6) E_{22} (Months)	(7) E_{23}
R&D Product	Light Trailer	3	B′	unknown	11	B′	2	A′
	Heavy Trailer	3	B′	unknown	20	B′	0	A’
	MPV	8	B′	unknown	7.3	B′	0	A′
	SRV	10	B′	unknown	9.7	B′	0	A′

Table 11. The Original assessment of R&D project in qualitative attributes

Qualitative Attributes		(1) e^3_{11}				(2) e^1_{12}					(3) e^2_{12}					(4) e^1_{13}					(5) e^2_{13}			
		A′	B′	C′	D′	A′	B′	C′	D′	E′	A′	B′	C′	D′	E′	A′	B′	C′	D′	E′	A′	B′	C′	D′
R&D Product	Light Trailer	4	3		2		6	3			5	2	2			1	4	3	1			1	4	4
	Heavy Trailer	10	2			8	4				3	8	1			5	6		1			3	7	2
	MPV	2	8			4	5	1			3	7				5	5				1	2	5	2
	SRV	7	3			6	3	1			1	4	5			8	2					2	2	6

Qualitative Attributes		(6) E_{21}			(7) e^1_{31}				(8) e^2_{31}					(9) e^3_{31}					(10) E_{32}			**Total number of experts**
		A′	B′	C′	A′	B′	C′	D′	A′	B′	C′	D′	E′	A′	B′	C′	D′	E′	A′	B′	C′	
R&D Product	Light Trailer	4	5		2	2	4	1	1	3	5			2	1	5	1		6	2	1	9
	Heavy Trailer	10	2		4	5	3		3	7	2			4	7	1			11		1	12
	MPV	5	4	1	2	4	4		3	5	2			6	3	1			7	2	1	10
	SRV	5	5		2	4	4		1	5	4			4	4	2			7	3		10

Table 12. The standard table for checkout and assessment system of strategic R&D project

General attributes	Criteria in the second level	Factors in the lowest level (contents of assessment)	Type of project			
			Light Trailer	Heavy Trailer	MPV	SRV
Quality of product $E_1(\omega_1)$	Scale and importance of project $E_{11}(\omega_{11})$	workload $e_{11}^1(\omega_{111})$	3	3	8	10
		origin of person $e_{11}^2(\omega_{112})$	$B'(1.0)$	$B'(1.0)$	$B'(1.0)$	$B'(1.0)$
		importance of project $e_{11}^3(\omega_{113})$	$A'(0.45)$ $B'(0.33)$ $D'(0.22)$	$A'(0.83)$ $B'(0.17)$	$A'(0.2)$ $B'(0.8)$	$A'(0.7)$ $B'(0.3)$
	Content of technique $E_{12}(\omega_{12})$	advance of critical techniques $e_{12}^1(\omega_{121})$	$B'(0.67)$ $C'(0.33)$	$A'(0.67)$ $B'(0.33)$	$A'(0.4)$ $B'(0.5)$ $C'(0.1)$	$A'(0.6)$ $B'(0.3)$ $C'(0.1)$
		ratio between quality and price $e_{12}^2(\omega_{122})$	$A'(0.56)$ $B'(0.22)$ $C'(0.22)$	$A'(0.25)$ $B'(0.67)$ $C'(0.08)$	$A'(0.3)$ $B'(0.7)$	$A'(0.1)$ $B'(0.4)$ $C'(0.5)$
		reliability of product $e_{12}^3(\omega_{123})$				
		economy $e_{12}^4(\omega_{124})$	11	20	7.3	9.7
	Theoretical value and level of innovation $E_{13}(\omega_{13})$	theoretical standard of project $e_{13}^1(\omega_{131})$	$A'(0.11)$ $B'(0.45)$ $C'(0.33)$ $D'(0.11)$	$A'(0.42)$ $B'(0.50)$ $D'(0.08)$	$A'(0.5)$ $B'(0.5)$	$A'(0.8)$ $B'(0.2)$
		degree of innovation $e_{13}^2(\omega_{132})$	$B'(0.10)$ $C'(0.45)$ $D'(0.45)$	$B'(0.25)$ $C'(0.58)$ $D'(0.17)$	$A'(0.1)$ $B'(0.2)$ $C'(0.5)$ $D'(0.2)$	$B'(0.2)$ $C'(0.2)$ $D'(0.6)$
		ratio of individual design $e_{13}^3(\omega_{133})$	$B'(1.0)$	$B'(1.0)$	$B'(1.0)$	$B'(1.0)$
Process control $E_2(\omega_2)$	Quality of project $E_{21}(\omega_{21})$		$A'(0.44)$ $B'(0.56)$	$A'(0.83)$ $B'(0.17)$	$A'(0.5)$ $B'(0.4)$ $C'(0.1)$	$A'(0.5)$ $B'(0.5)$
	Completion time for a project $E_{22}(\omega_{22})$		2 months	0	0	0
	Investment $E_{23}(\omega_{23})$		$A'(1.0)$	$A'(1.0)$	$A'(1.0)$	$A'(1.0)$
Added value by project $E_3(\omega_3)$	Project team $E_{31}(\omega_{31})$	documents of rules and regulations established about project group $e_{31}^1(\omega_{311})$	$A'(0.22)$ $B'(0.22)$ $C'(0.45)$ $D'(0.11)$	$A'(0.33)$ $B'(0.42)$ $C'(0.25)$	$A'(0.2)$ $B'(0.4)$ $C'(0.4)$	$A'(0.2)$ $B'(0.4)$ $C'(0.4)$
		the routine operational management documents $e_{31}^2(\omega_{312})$	$A'(0.11)$ $B'(0.33)$ $C'(0.56)$	$A'(0.25)$ $B'(0.58)$ $C'(0.17)$	$A'(0.3)$ $B'(0.5)$ $C'(0.2)$	$A'(0.1)$ $B'(0.5)$ $C'(0.4)$
		management documents about R&D process of products $e_{31}^3(\omega_{313})$	$A'(0.22)$ $B'(0.11)$ $C'(0.56)$ $D'(0.11)$	$A'(0.33)$ $B'(0.58)$ $C'(0.09)$	$A'(0.6)$ $B'(0.3)$ $C'(0.1)$	$A'(0.4)$ $B'(0.4)$ $C'(0.2)$
	Continuity of technique $E_{32}(\omega_{32})$	accumulation and continuity of technique ω_{32}	$A'(0.67)$ $B'(0.22)$ $C'(0.11)$	$A'(0.92)$ $C'(0.08)$	$A'(0.7)$ $B'(0.2)$ $C'(0.1)$	$A'(0.7)$ $B'(0.3)$

References

1. G. Balestra and A. Tsoukias, “Multicriteria analysis represented by artificial intelligence techniques”, *J. Oper. Res. Soc.*, **41** (5), 419–430 (1990).
2. A. Bufardi, “On the construction of fuzzy preference structures”, *J. Multi-Criteria Decision Analysis*, **7** (3), 57–88 (1998).
3. A.P. Dempster, “Upper and lower probabilities induced by a multivalued mapping”, *Ann. of Math. Statis.*, **38**, 325–339 (1967).
4. C.L. Huang and K. Yoon, *Multiple Attribute Decision Making Methods and Applications, A State-of-Art Survey*, New York, Springer-Verlag, 1981.
5. R.L. Keeney and H. Raiffa, *Decisions With Multiple Objectives*, 2nd ed., Cambridge Univ. Press, Cambridge, U.K., 1993.
6. X.B. Liu, S.L. Yang, “Hadamard convex set of judgmental matrices”, *System Engineering Theory and Practice*, **4**, 83–85 (2000).
7. G. Shafer, *A mathematical theory of evidence*, Princeton University Press, Princeton, 1976.
8. T.L. Saaty, *The Analytic Hierarchy Process*, Pittsburgh, PA, Univ. Pittsburgh, 1988.
9. T.L. Saaty, “How to make a decision: The Analytic Hierarchy Process”, *Europ. J. of Operational Research*, **48** (1), 9–26 (1990).
10. P. Sen and J.B. Yang, “Multiple criteria decision making in design selection and synthesis”, *J. Eng. Design*, **6** (3), 207–230 (1995).
11. M. Snmez, G.D. Holt, J.B. Yang and G. Graham, “Applying evidential reasoning to pre-qualifying construction contractors”, *J. of Management in Engineering*, **18** (3), (2002).
12. J. Wang, J.B. Yang, and P. Sen, “Safety analysis and synthesis using fuzzy sets and evidential reasoning”, *Reliability Engineering and System Safety*, **47**, 103–118 (1995).
13. J. Wang and J.B. Yang, “A subjective safety based decision making approach for evaluation of safety requirements specifications in software development”, *Int. J. Reliab., Qual., Saf. Eng.*, **8** (1), 35–57 (2001).
14. Y.M. Wang, J.B. Yang, D.L. Xu, “A preference aggregation method through the estimation of utility intervals”, *Computers & Operations Research*, **32** (8), 2027–2049 (2005).
15. Y.M. Wang, J.B. Yang, and D.L. Xu, “Environmental impact assessment using the evidential reasoning approach”, *Europ. J. of Operational Research*, **174**, 1885–1913 (2006).
16. C.C. White, “A survey on the integration of decision analysis and expert systems for decision support”, *IEEE Trans. Syst., Man, Cybern.*, **20** (2), 358–364 (1990).
17. D.L. Xu, J.B. Yang, Y.M. Wang, “The evidential reasoning approach for multi-attribute decision analysis under interval uncertainties”, *Europ. J. of Operational Research*, **174**, 1914–1943 (2006).
18. R.R. Yager, “Decision-making under various types of uncertainties”, *J. Intell. Fuzzy Syst.*, **3** (4), 317–323 (1995).
19. J.B. Yang and M.G. Singh, “An evidential reasoning approach for multiple attribute decision making with uncertainty”, *IEEE Trans. Syst., Man, Cybern*, **24** (1), 1–18 (1994).
20. J.B. Yang and P. Sen, “A general multi-level evaluation process for hybrid MADM with uncertainty”, *IEEE Trans. Syst., Man., Cybern.*, **24** (10), 1458–1473 (1994).
21. J.B. Yang and P. Sen, “Multiple attribute design evaluation of large engineering products using the evidential reasoning approach”, *J. Eng. Design*, **8** (3), 211–230 (1997).
22. J.B. Yang, “Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainties”, *Europ. J. of Operational Research*, **131**, 31–61 (2001).
23. J.B. Yang, B.G. Dale, and H.R. Siow, “Self-assessment of excellence: an application of the evidential reasoning approach”, *Int. J. of Production Research*, **39** (16), 3789–3812 (2001).
24. J.B. Yang and D.L. Xu, “On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty”, *IEEE Trans. Syst., Man., Cybern., – Part A: Systems and Humans*, **32** (3), 289–304 (2002).
25. J.B. Yang, Y.M. Wang, D.L. Xu, and K.S. Chin, “The evidential reasoning approach for MCDA under both probabilistic and fuzzy uncertainties”, *Europ. J. of Operational Research*, **171** (1), 309–343 (2006).
26. Z. Jun, “The indirect presentation of judgmental matrix in AHP”, *Systems Engineering*, **6** (1988).
27. B. Roy, “The problems and methods with multiple objective functions”, *Mathematical Programming*, **1**, 239 (1971).
28. B. Roy and P. Bertier, “La method ELECTRE II: Une application au media-planning”, in: M. Ross (Ed.), *Operational Research 1972*, North-Holland Publishing Company, 291–302 (1973).
29. J.P. Brans, B. Mareschal, and Ph. Vincke, “Promethee: A new family of outranking methods in multicriteria analysis”, in: J.P. Brans (Ed.), *Operational Research’84*, North-Holland, Elsevier Science Publishers, 408–421 (1984).
30. R. Whiteley, T. Parish, and R. Dressler, “Evaluating R&D Performance Using the New Sales Ratio”, *Research Technology Management*, **41** (5), 20–22 (1998).
31. Szakony i, Hobert, “Measuring R&D Effectiveness-I”, *Research-Technology Management*, **37** (2) (1994).
32. Szakony i, Hobert, “Measuring R&D Effectiveness-

II", *Research-Technology Management*, **37** (3) (1994).

33. R. Dressler, R.S. Wood, and V. Alvarez, "Evaluating R&D Performance Using the Cost Savings Metric", *Research Technology Management*, **42** (2), 13–14 (1999).
34. Dronggelen, Inge C. Kerssens-van & Cook, Andrew, "Design principles for the development of measurement for research and development processes", *R&D Management*, **27** (4), 345–357 (1997).
35. Y.-II Song, D.-H. Lee, Y.-G. Lee, Y.-C. Chung, "Managing uncertainty and ambiguity in frontier R&D projects: A Korean case study", *J. Eng. Technol. Manage.*, **24**, 231–250 (2007).
36. W.R. Dincan, "A guide to the project management body of knowledge", PMI, 1996.
37. De hai, Wang, "Theory and method of modern project management", China Agriculture Press, 1998.