INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, 13(3), 391-407, June 2018.

Attribute Selection Method based on Objective Data and Subjective Preferences in MCDM

X. Ma, Y. Feng, Y. Qu, Y. Yu

Xiaofei Ma, Yi Feng
Technology Planning
Dalian Commodity Exchange
Dalian City, China, 116023
maxf@dce.com.cn, fengyi@dce.com.cn

Yi Qu*, Yang Yu
Agricultural Bank of China Data Center
88 Aoni Road, Pudong New Area
Shanghai City, China, 200131
*Corresponding author: 57638907@qq.com
yuyinyang@126.com

Abstract: Decision attributes are important parameters when choosing an alternative in a multiple criteria decision-making (MCDM) problem. In order to select the optimal set of decision attributes, an analysis framework is proposed to illustrate the attribute selection problem. Then a two-step attribute selection procedure is presented based on the framework: In the first step, attributes are filtered by using correlation algorithm. In the second step, a multi-objective optimization model is constructed to screen attributes from the results of the first step. Finally, a case study is given to illustrate and verify this method. The advantage of this method is that both external attribute data and subjective decision preferences are utilized in a sequential procedure. It enhances the reliability of decision attributes and matches the actual decision-making scenarios better.
Keywords: attribute selection, multi-criteria decision-making (MCDM), multi-objective optimization, attribute correlation.

1 Introduction

Multi-criteria decision-making (MCDM) is successfully applied to help decision makers (DMs) choose optimal alternatives. During the past few decades, various MCDM methods have been proposed based upon different philosophies such as multi-attribute utility, the analytic hierarchy process (AHP), outranking methods and so on. Meanwhile, many decision support systems (DSSs) have been designed in MCDM to assist DMs in analyzing problems and making decisions more easily. MCDM deals with a general class of problems that contains multiple attributes, objectives and criteria [20]; and alternatives are determined based on many qualitative or quantitative criteria which are generally complicated and assessed by more relevant attributes. However, not all of them can be used in decision-making procedure, because they contain plenty of redundant or "noisy" attributes and it will lead to useless decision alternatives. Therefore, the effectiveness of an alternative is highly dependent on the set of decision attributes.

The concept of "optimal" can be illustrated by two aspects: (1) Elements of the set are highly related to the MCDM problem for decision purpose; (2) The set of attributes is parsimonious, and the selected alternative will be suboptimum if one of these attributes is omitted. The rationale of attribute selection is similar to feature selection in data mining field. There are a lot of feature selection methods have been proposed, but only a few of them can be applied in decision-making process directly which are mentioned in the Literature Review part. Most of these methods can't account for both objective and subjective perspectives, so there is a low reliability of decision attributes. This study is focused on how to select the optimal set of decision attributes for a MCDM problem, external attribute data and subjective decision preferences are utilized in a sequential procedure to enhance the reliability of decision attributes.

The attribute selection method for decision problem should account for both objective and subjective perspectives, more precisely, this paper merges attribute data (objective) with DMs' requirement (subjective). Utilize the former for rational, constrained modeling and the latter for adapting specific problem issues to the decision-making process. In order to obtain the optimal set of attributes, an analysis framework for attribute selection problem and a two-step screening procedure is conducted in this paper. The specific procedure is illustrated in Figure 1.

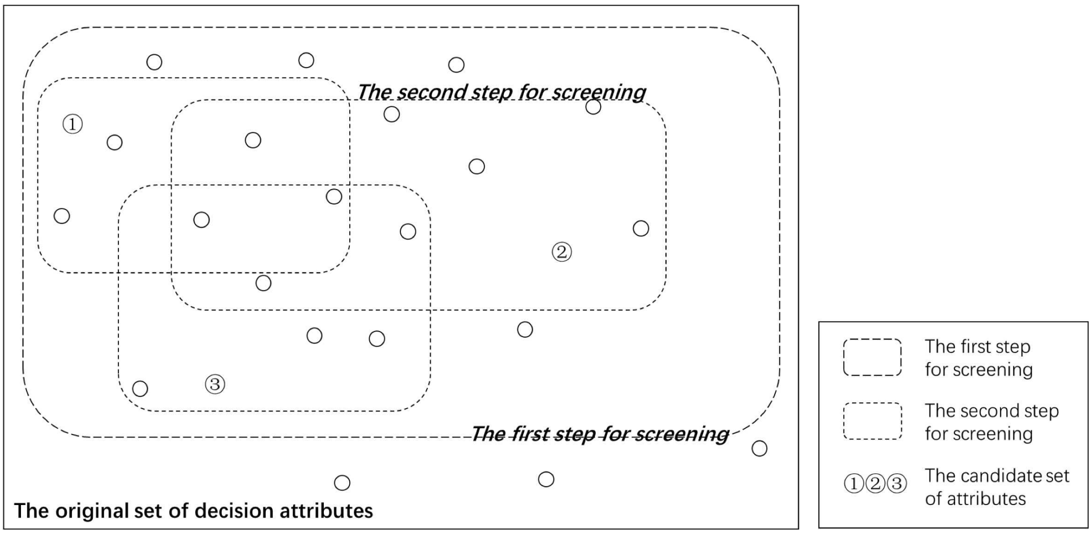

Figure 1: The illustration for the presented method

This new attribute selection method in MCDM contributes to selecting the optimal set of decision attributes and helping DMs choose optimal alternatives better. Meanwhile, this method merges objective data and subjective preferences to improve performance for actual decision scenarios.

2 Literature review

Many feature selection methods have been proposed, and they are usually classified into three classes: "filter" methods, "wrapper" methods and "embedded" methods [2,7,8,14]. Wu [19] proposed using the fuzzy and grey Delphi methods to identify a set of reliable attributes and, based on these attributes, transforming big data to a manageable scale to consider their impacts. Meinshausen et al. [11] demonstrated linear model and Gaussian model of variable selection consistency in higher dimensional case. Although they have good performance, they cannot be applied in decision-making process directly (partial principle may be accepted). Only a few papers mentioned attribute selection in decision-making problems; for instance, Chun [4] considered the "optimizing" and "satisficing" (a portmanteau of satisfy and suffice) attributes and deal with the multi-attribute decision problem with sequentially presented decision alternatives; Dai et al. [6] constructed three attribute selection approaches in context of incomplete decision systems based on information-theoretical measurement of attribute importance; in order to screen the critical factors influencing the stability of perilous rock, Meng et al. [12] used fuzzy compromise TOPSIS method to calculate the importance of attributes in the decision problem.

Wu [18] used Fuzzy Delphi method to screen out the unnecessary attributes to deal with the complex interrelationships among the aspects and attributes. Attribute selection procedure can improve the decision performance, but no papers used this procedure in MCDM problems.

3 Method

3.1 Preliminaries

Trapezoidal fuzzy numbers

Definition 1. Let X be a universe set. A fuzzy $\widetilde{a}$ in a universe of discourse X is characterized by a membership function $\mu_{\widetilde{a}}(x)$, which associates with each element x in X, a real number in the interval [0,1]. The function is termed the grade of membership of x in $\widetilde{a}$.

Definition 2. A tuple $\widetilde{A} = (a, b, c, d), a \leq b \leq c \leq d$, is called a trapezoidal fuzzy number (TFN) if its membership function is

$$\mu_{\widetilde{a}}(x) = \begin{cases} (x-a)/(b-a) & a \leq x \leq b \\ 1 & b \leq x \leq c \\ (d-x)/(d-c) & c \leq x \leq d \\ 0 & otherwise \end{cases}$$

Where $a,\ b,\ c,\ d$ are real numbers.

Definition 3. Given two trapezoidal fuzzy numbers $\widetilde{A} = (a_1, b_1, c_1, d_1)$, $\widetilde{A} = (a_2, b_2, c_2, d_2)$, and a real number λ, the main operations can be expressed as follows:
①$\widetilde{A_1} \bigoplus \widetilde{A_2} = (a_1 + a_2, b_1 + b_2, c_1 + c_2, d_1 + d_2)$
②$\widetilde{A_1} \bigotimes \widetilde{A_2} = (a_1 a_2, b_1 b_2, c_1 c_2, d_1 d_2)$
③$\lambda \bigotimes \widetilde{A_2} = (\lambda a_1, \lambda b_1, \lambda c_1, \lambda d_1)$

Multi-objective optimization

Let χ be a vector containing n decision variables and in a universe of discourse X. Mathematically, an optimization problem with p objective functions can be expressed as (Mahdi and Seyed, 2012):

minimize $f_i(x)$ for $i = 1, 2, ..., p$
subject to: $g_j(x) \leq 0, j = 1, 2, ..., q,$
$h_j(x) = 0, j = q + 1, 2, ..., m$

where $x = (x_1, x_2, ..., x_n)$, x_l is the lth decision variable; p and m are respectively the numbers of objective function and constraint.

A variety of methods can be used to solve this problem. One popular method is to combine those objectives into a single composite objective so that traditional mathematical programming methods can be applied. And other approaches are based on the Pareto optimum concept, and more specifically, let $y = (y_1, y_2, ..., y_n)$ be another vector containing **n** decision variables in X.

Definition 4. (Domination) $x \in X$ dominates $y \in X$, denoted $x \succ y$, if $\forall j \in p$, $v_j(x) \leq v_j(y)$ with at least one strict inequality.

Definition 5. (Pareto Optimal) x is a Pareto Optimal (PO) solution, if there is no $y \in X$ such that $y \succ x$. The set of all Pareto optimal solutions in X is denoted PO(X).

These methods include Multiple Objective Genetic Algorithm (MOGA), Non-dominated Sorting Genetic Algorithm (NSGA), NSGA-II, Multi-objective Multi-state Genetic Algorithm (MOMS-GA), Niched-Pareto Genetic Algorithm (NPGA) and Multi-objective Particle Swarm Optimization (MOPSO) [9, 21].

An analysis framework of attribute selection under MCDM

Multi-Criteria Decision Making (MCDM) has been widely applied as a well-known branch of decision making, and can be further divided into multi-objective decision making (MADM) and multi-attribute decision making (MODM) [5, 13]. The whole hierarchical structure of a MCDM problem is shown in Figure 2, and it can be unfolded from four hierarchies from top to the bottom, namely, objective, criteria, attribute and alternative.

(1) Objective Hierarchy: Generally contain multiple decision objectives given by DMs which reflect different decision requirement.

(2) Criteria Hierarchy: Criteria Hierarchy is to realize a fair comparison of alternatives from the various aspects of Objective Hierarchy, and it is also developed in a hierarchical fashion, starting from some general but imprecise criteria description, which is refined into more precise sub- and sub-sub criteria.

(3) Attribute Hierarchy: Attribute Hierarchy is determined by criteria Hierarchy and Objective Hierarchy. It is generally assumed that each criterion can be represented by some measurable attributes of the consequences arising from implementation of any particular decision alternative [16], as well as objective. No matter what kind of MCDM methods, decision attributes are the bottom and direct evaluative parameters to determine alternatives.

(4) Alternative Hierarchy: Consist from candidate alternatives evaluated by the set of decision attribute.

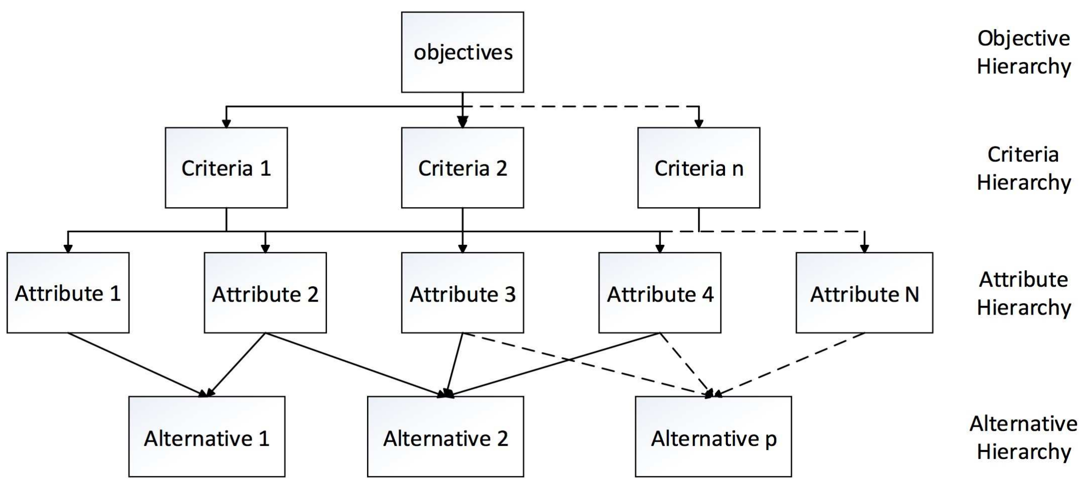

Figure 2: The hierarchical structure of a MCDM problem

This paper focus on how to obtain the optimal set of decision attributes from Attribute Hierarchy for evaluating alternatives.

Rationale of the proposed method

Let $O = \{O^1, O^2, ..., O^\gamma\}$ be a finite set of objectives, and $C = \{C_1, C_2, ..., C_n\}$ be the set of criteria. In most actual cases, objectives and criteria are always predetermined by MDs according to different knowledge background and decision experiences, thus we assume that all of objectives and criteria (or sub-criteria) are given.

We define decision attribute selection procedure as a sequential selection procedure based on objective and subjective information. It means that a series of screening steps will be applied in sequence to obtain an optimal set of attributes. A sequential screening can be described as follows [3]:

$$Scr_{h,h-1,...,1}(A_{original}) = Scr_h(Scr_{h-1}(...(Scr_1(A_{original}))...)) \tag{1}$$

where $Scr_k, k = 1, 2, ..., h$ are screening steps, and Scr_h is the final screening step in this sequential screening; $A_{original} = \{a_j | j \in N\}$ is the original set of decision attributes.

And two screening steps are conducted in proposed selection procedure, namely, $Scr_{2,1}(A) = Scr_2(Scr_1(A))$.

Scr_i is a screening step based on real-time attribute data, and the corresponding screening algorithm is established in the step. The result $Scr_i(A)$ is to obtain a relatively optimized set of decision attributes A_{Scr_1} by deleting redundant attributes.

Scr_2 is a further screening step based on the results of $Scr_2(A)$ to acquire the absolutely optimal set of decision attributes A_{Scr_2}. Decision goals that represent subjective preference are utilized in this step, and the selected set should satisfy all decision goals.

The whole procedure of attribute selection is shown by Figure 3

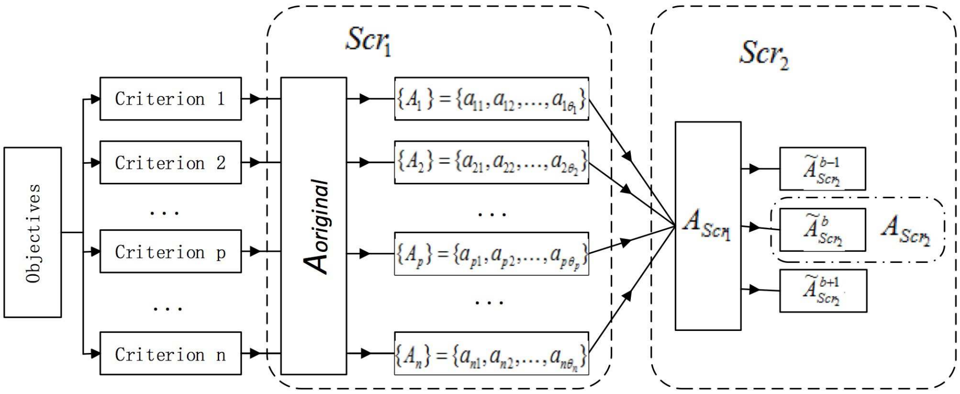

Figure 3: The procedure of the proposed method

where $A_{Scr_1} = \{\{A_1\}, \{A_1\}, ..., \{A_p\}, ..., \{A_n\}\}$, and $\widetilde{A}^b_{Scr_2}$ is candidate optimal set of Scr_2.

3.2 A new method for attribute selection

Preprocessing of the data

We consider three data types of decision attributes in this paper, namely, real numbers expressed as α^0, intervals expressed as $[\alpha^L, \alpha^U]$ and linguistic term set expressed as a linguistic term set $S = \{s_0, s_1, ..., s_K\}$. And real-time attribute data can be listed as attribute value series in Table 1.

where $c \le t_1 < t_2 < ... < t_m \le d$ and $[c, d]$ is a time interval.

Table 1: Attribute value series

	α_1	α_2	...	α_j	...
t_1	$\alpha_1(t_1)$	$\alpha_2(t_1)$	...	$\alpha_j(t_1)$	...
t_2	$\alpha_2(t_2)$	$\alpha_2(t_2)$	...	$\alpha_j(t_2)$	...
...					
t_i	$\alpha_i(t_i)$	$\alpha_2(t_i)$	...	$\alpha_j(t_i)$	...
...					
t_m	$\alpha_m(t_m)$	$\alpha_2(t_m)$	...	$\alpha_j(t_m)$	...

Both Scr_1 and Scr_2 are quantification procedures and therefore, in order to guarantee the scientificity and preciseness of the presented method, preprocessing of attribute data need to be carried out firstly (referring to our previous work [10]).

(1) Convert attribute values to TFNs

For a real number α^0, it can be denoted as a TFN $(\alpha^1, \alpha^2, \alpha^3, \alpha^4)$, where $\alpha^1 = \alpha^2 = \alpha^3 = \alpha^4 = \alpha^0$; for internal value $\lfloor \alpha^L, \alpha^U \rfloor$, the corresponding TFN can also be expressed as $(\alpha^1, \alpha^2, \alpha^3, \alpha^4)$, where $\alpha^1 = \alpha^2 = \alpha^L$ and $\alpha^3 = \alpha^4 = \alpha^U$. It is a little complex for a linguistic term set with odd cardinality $S = \{s_0, s_1, ..., s_K\}$, and the element s_K is converted to the corresponding TFN as follows:

$$(\alpha_k^1, \alpha_k^2, \alpha_k^3, \alpha_k^4) = \left(\max\left\{\frac{2k-1}{2K+1}, 0\right\}, \frac{2k}{2K+1}, \frac{2k+1}{2K+1}, \min\left\{\frac{2k+2}{2K+1}, 1\right\}\right) \tag{2}$$

where $k = 0, 1, ..., K$.

(2) Normalize values of attributes

In real-world decision scenarios, some decision attributes are cost ones, which mean the lower the values of attributes, the better they will be; the others are benefit ones, which mean the higher the values of attributes, the better they will be. Without loss of generality, we assume that the attribute values can be expressed as $\alpha_j(t_i) = (\alpha_j^1(t_i), \alpha_j^2(t_i), \alpha_j^3(t_i), \alpha_j^4(t_i))$, and normalized attribute values can be denoted as $\gamma_j(t_i) = (r_j^1(t_i), r_j^2(t_i), r_j^3(t_i), r_j^4(t_i))$. The normalized methods are given as follows:

- For cost attributes:

$$r_j^v(t_i) = \frac{\max\limits_i\left(\alpha_j^4(t_i)\right) - \alpha_j^v(t_i)}{\max\limits_i\left(\alpha_j^4(t_i)\right) - \min\limits_i\left(\alpha_j^1(t_i)\right)} \quad v = 1, 2, 3, 4. \tag{3}$$

- For benefit attributes:

$$r_j^v(t_i) = \frac{\alpha_j^v(t_i) - \min\limits_i\left(\alpha_j^1(t_i)\right)}{\max\limits_i\left(\alpha_j^4(t_i)\right) - \min\limits_i\left(\alpha_j^1(t_i)\right)} \quad v = 1, 2, 3, 4. \tag{4}$$

And the normalized attribute value series can be denoted as
$\gamma_1 = \{r_1(t_1), r_1(t_2), ..., r_1(t_m)\}$;

$\gamma_2 = \{r_2(t_1), r_2(t_2), ..., r_2(t_m)\}$;
...;
$\gamma_j = \{r_j(t_1), r_j(t_2), ..., r_j(t_m)\}$;
....

The first step for screening Scr_1

In Scr_1, one "relatively best" representative attribute will be chosen by DMs for each criterion. It means that,

For $\forall C_p \in \{C_1, C_2, ..., C_n\}, 1 \leq p \leq n$,

$\exists a_{p1} \in A_{original}, st.C_p \Leftrightarrow a_{p1}$.

For example, "Economic losses" is a criterion in emergency decision, and "Amount of loss" is the "relatively best" attribute to evaluate loss degree. Thus the set $\{a_{p1}|1 \leq p \leq n\}$ is a significant by-product after criteria determination, and actually it is the input of Scr_1.

Scr_1 is used to screen decision-relevant attributes from $A_{original}$ by analyzing attribute interrelation. More precisely, we will investigate geometrical relationship between attributes based on attribute value series; select highly correlative attributes for a_{p1} from $A_{original}$ to constitute a "relatively best" representative set for each criterion C_i, and omit the rest of decision-irrelevant attributes. The step is entirely implemented on attribute value series without subjective impacts, and the outcome A_{Scr_1} is a smaller and relative optimal set which the elements are determined by evaluation criteria.

The idea of Scr_1 arises from Grey Relational Analysis (GRA) Theory which provides a valid way to quantify the interrelation of different factors using geometric methods [22]. And the detailed process of Scr_1 is given in the following.

Step 1.1: *Normalize attribute value series*

It is different from normalization of attribute values, and it guarantees comparability of different attribute value series. The specific process is

$$y_j = \{a_j(t_i)/D_j, i = 1, 2, ..., m\} \tag{5}$$

$$D_j = \frac{1}{m-1} \sum_{i=2}^{m} |a_j(t_i) - a_j(t_{i-1})| \tag{6}$$

where y_j is normalized attribute value series, and D_j is increment average of y_j.

Step 1.2: *Calculate increment series*

$$\Delta y_j = \{\Delta y_j(t_i) = y_j(t_i) - y_j(t_{i-1}), i = 2, 3, ..., m\} \tag{7}$$

Step 1.3: *Calculate correlation coefficient*

A_{Scr_1} is obtained based on the input set $\{a_{p1}|1 \leq p \leq n\}$ and thus, Step 1.3 will be stepwise conducted for each element a_{p1}. The correlation coefficient of different t_i between a_{p1} and a_j can be defined as follows:

$$\xi(t_i) = \begin{cases} sgn(\Delta y_{p1}(t_i) \cdot \Delta y_j(t_i)) \cdot \frac{\min(|\Delta y_{p1}(t_i)|, |\Delta y_j(t_i)|)}{\max(|\Delta y_{p1}(t_i)|, |\Delta y_j(t_i)|)}; & j \neq p1 \\ 0 & \Delta y_{p1}(t_i) \cdot \Delta y_j(t_i) = 0 \end{cases} \tag{8}$$

where $sgn(\cdot)$ is a sign function; and $sgn(\Delta y_{p1}(t_i) \cdot \Delta y_j(t_i)) = 1$, if $\Delta y_{p1}(t_i) \cdot \Delta y_j(t_i) > 0$; $sgn(\Delta y_{p1}(t_i) \cdot \Delta y_j(t_i)) = -1$, if $\Delta y_{p1}(t_i) \cdot \Delta y_j(t_i) < 0$.

Step 1.4: *Calculate correlation degrees*

The correlation degree between a_{p1} and a_j can be obtained by

$$\tau(a_{p1}, a_j) = \left| \frac{1}{d-c} \sum_{i=2}^{m} \Delta t_i \cdot \xi(t_i) \right| \tag{9}$$

The correlation degrees reflect the closeness of two attributes; the greater the value of $\tau(a_{p1}, a_j)$ is, the closer the two attributes are, and the better the capability of a_j is to evaluate criterion C_p.

Step 1.5: *Set threshold and screen correlation attributes for a_{p1}*

For $\forall C_p$, a threshold of correlation degree τ^* should be set by DMs, and the attribute a_j will be reserved, if $\tau(a_{p1}, a_j) > \tau^*$. And the reserved attributes for evaluating C_p can be denoted as $\{A_p\} = \{a_{p1}, a_{p2}, ..., a_{p\theta_p}\}$, where θ_p is the sum of attributes. In addition, the correlation degree $\tau(a_{p1}, a_{pq_p})$ between a_{p1} and a_{pq_p} will be abbreviated as z_{pq_p}, where $1 \leq q_p \leq \theta_p$.

Step 1.6: *Acquire the screening result set A_{Scr_1}*

Repeat Step 1.1 to Step 1.5 for each a_{p1}, where $1 \leq p \leq n$; and the result set A_{Scr_1} can be obtained, namely,

$A_{Scr_1} = \{\{A_1\}, \{A_n\}, ..., \{A_n\}\}$.

The second step based on multi-objective optimization Scr_2

Implementing decision attribute selection is aimed to assist DMs in evaluating alternatives for given decision problems and therefore, the selected attributes should be sensitive to the decision problem as well as familiar to DMs. Sensitivity reflects whether the selected attributes are accurate enough to represent the decision problem, and familiarity is used to estimate whether DMs are certain enough to make decision based on the selected attributes. For example, "fire duration" and "indoor fire load" are more important (sensitive) attributes than "carbon monoxide concentration" for evaluating fire grades; while it is more reliable for medical staff to estimate the physical condition of the wounded with "carbon monoxide concentration" than "Radiative Heat Flux".

Multi-objective optimization theory is a feasible way to quantify and synthesize these requirements, and it has solved similar problems in other areas successfully [1]. Thus three main objectives will be considered in this paper: attribute amount, attribute utility and attribute familiarity. Attribute amount should be as few as possible without reducing the validity of alternative, and it is the demand of both Scr_2 and attribute selection method. Attribute utility is used to distinguish attribute sensitivity in estimating the same decision problem, and attribute familiarity explores DMs' cognition differences for decision attributes. The outcome A_{Scr_2} drawn from Scr_2 should satisfy mentioned three goals, and now the optimal set of decision attributes have been found after Scr_1 and Scr_2, namely, $A_{Scr_2} = Scr_2(Scr_1(A_{original}))$.

A set of zero-one binary vectors $\{v^b\}$ with $T = (\theta_1+\theta_2+...+\theta_n)$ dimensionality will be defined to represent different candidate outcome set $\widetilde{A}^b_{Scr_1}$, where one represents the attribute belonging to the set, otherwise it is zero. The optimal set of decision attributes is $A_{Scr_2} \in \{\widetilde{A}^b_{Scr_1} | b \in N\}$, and $\{\widetilde{A}^b_{Scr_1}\}$ contains all the subsets of A_{Scr_1}.

Three objective functions are defined firstly, and we will show that all the objective functions can be expressed as the functions with independent variable v^b.

Defining objective functions (1) Attribute utility function

Attribute utility function measures decision sensitivity of $\widetilde{A}^b_{Scr_1}$, and it can be defined as follows:

$$U(\widetilde{A}^b_{Scr_1}) = g(MU(\widetilde{A}^b_{Scr_1}), W(\widetilde{A}^b_{Scr_1})) \tag{10}$$

where

$$\text{①}\ MU(\widetilde{A}^b_{Scr_1}) = v^b \cdot MU^* \tag{11}$$

MU^* represents attribute utility matrix, and it is diagonal matrix with the diagonal elements $\frac{\Delta a_{11}}{\Delta a_{11}}, ..., \frac{\Delta a_{11}}{\Delta a_{1\theta_1}}, ..., \frac{\Delta a_{p1}}{\Delta a_{pq_p}}, ..., \frac{\Delta a_{n1}}{\Delta a_{n1}}, ..., \frac{\Delta a_{n1}}{\Delta a_{n\theta_n}}$.

More precisely, MU^* can be explained by a marginal utility function $u(\cdot)$, namely,

$$u(a_{pq_p}) = \frac{\Delta a_{p1}}{\Delta a_{pq_p}} \qquad 1 \leq p \leq n, 1 \leq q_p \leq \theta_p; \tag{12}$$

where Δa_{p1} is increment of a_{p1}, and Δa_{pq_p} is the corresponding varying of a_{pq_p}. And the function $u(\cdot)$ quantifies the assessment performance on criterion C_i of the attribute a_{pq_p} according to calculate the ratio between increments of the two attribute values.

In fact, MU^* can be predetermined by domain experts, because plenty of field studies have been conducted to investigate utility relation of attributes.

$$\text{②}\ W(\widetilde{A}^b_{Scr_1}) = W^* \cdot (v^b)' \tag{13}$$

W^* represents attribute weighting matrix, and it is also a diagonal matrix with diagonal elements $w_{11}, ..., w_{1\theta_1}, ..., w_{pq_p}, ..., w_{n1}, ..., w_{n\theta_n}$.

where

$$w_{pq_p} = \omega_p \cdot z_{pq_p}, \quad 1 \leq p \leq n, 1 \leq q_p \leq \theta_p; \tag{14}$$

and $\{\omega_1, \omega_2, ..., \omega_n\}$ are the know weights of $C = \{C_1, C_2, ..., C_n\}$.

$g(\cdot)$ is a dot product function and therefore, $U(\widetilde{A}^b_{Scr_1})$ can be transformed into the following function with independent variable v^b, namely,

$$U(v^b) = (v^b \cdot MU^*) \cdot (W^* \cdot (v^b)') \tag{15}$$

(2) Attribute familiarity function

Attribute familiarity function is established to quantify DMs' cognition degrees on different candidate set $\widetilde{A}^b_{Scr_1}$, and it is defined as

$$F(v^b) = \eta \cdot \sigma \cdot (v^b)' \tag{16}$$

In Eq. 16, $\eta = (\eta_1, \eta_2, ..., \eta_L)_{1\times L}$ is the weight vector of DMs, and L is the number of DMs; σ is attribute familiarity matrix with the following expression

$$\sigma = \begin{pmatrix} \sigma_1(1) & \sigma_2(1) & ... & \sigma_T(1) \\ \sigma_1(2) & \sigma_2(2) & ... & \sigma_T(2) \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ \sigma_1(L) & \sigma_2(L) & ... & \sigma_T(L) \end{pmatrix}_{L\times T} \tag{17}$$

and $\sigma_h(l)$ represents the lth DM's familiarity degree with attribute a_h, where $1 \leq l \leq L$, and $1 \leq h \leq T$. Attribute familiarity is usually evaluated qualitatively by DMs, and many methods can be applied to quantify assessment results and obtain matrix σ, such as AHP and fuzzy set theory.

(3) Attribute count function

Attribute count function calculates the attribute number of a candidate set $\widetilde{A}^b_{Scr_1}$, and it can be defined by means of vector length formula,

$$Count(\widetilde{A}^b_{Scr_1}) = |v^b|^2 \tag{18}$$

where $|v^b|$ represents the length of vector v^b.

Establish multi-objective optimization model for Scr_2 After define relevant objective functions, the multi-objective optimization model of Scr_2 is established to obtain the Pareto Optimal solutions A_{Scr_2}, as well as the optimal set of decision attribute selection problem.

$$\begin{aligned} Maximize \quad & U(v^b) = (v^b \cdot MU^*) \cdot (W^* \cdot (v^b)') \\ & F(v^b) = \eta \cdot \sigma \cdot (v^b)' \end{aligned} \tag{19}$$

$$Minimize \quad Count(\widetilde{A}^b_{Scr_1}) = |v^b|^2$$

$$Subject\ to: \quad |v^b|^2 \geq n$$

The model Eq. 19 can be solved by aforementioned algorithm, such as MOGA, NSGA, MOMS-GA and traditional mathematical programming methods, and it depends the specific model drawn from given decision problems and decision requirements.

4 Empirical results

4.1 Accident scenario and computational settings

In order to validate the presented method, a fire accident scenario will be constructed and simulated. The fire happens in a factory dormitory, and two workers are asleep; some inflammable are stacked near the door which arise spontaneous combustion for some reasons. The total layout of this room is illustrated in Figure 4, containing three beds, two workers marked by fellow cuboids, fire origin and plenty of fire smoke. Now two experts need to make rescue alternative based on the collected data; one is fire expert who are engaged in fire research for many years and have in-depth study of the fire attributes, and the other is an experienced firefighter who know some medical knowledge. And they are weighted with 0.6 and 0.4 respectively. A fire simulation software FDS [17] is used to establish fire scenario and generate fire data.

The purpose of simulation is not only to provide a fire scenario, and more importantly, it can generate relevant fire attributes data for the following experiment. The data can be exported from FDS containing attribute names, physical dimensions, attribute values and attribute types. In this paper, we generate twelve fire attributes and 125 data records.

4.2 Select fire attributes by utilizing the presented method

In this fire scenario, three main criteria will be considered, i.e., fire behavior, personal security and environmental conditions, and corresponding attributes are shown in Table 5

where "Enthalpy" is an interval attribute, "Indoor fire load" and "Circuit aging degree" are linguistic attributes, and the rest are real number attributes. The bold attributes are "relatively best" representative attributes for each criterion.

(1) Data preprocessing

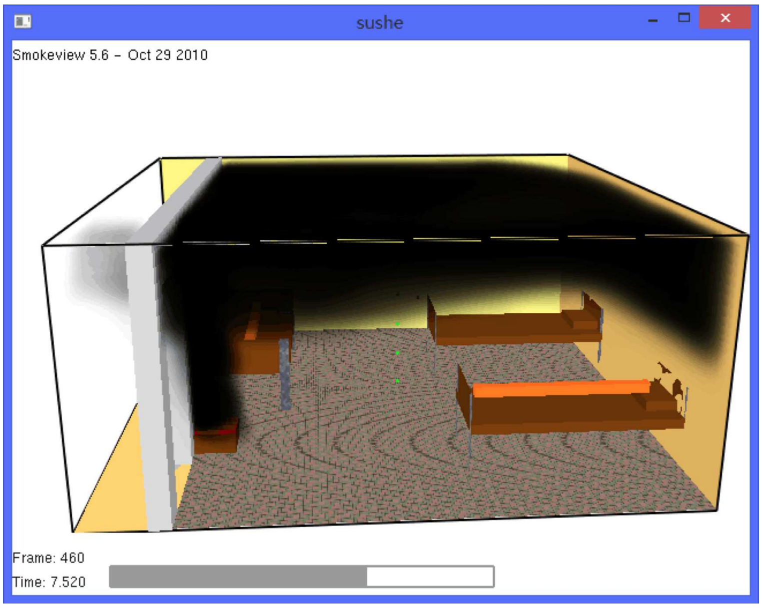

Figure 4: A simulated fire scenario

Table 2: List of criteria and corresponding attributes

Criterion	Attributes
Fire behavior	**Temperature** / Enthalpy / Heat Release Rate / Relative Humidity / Burning Rate / Pressure / Radiative Heat Flux
Personal security	**Carbon monoxide** / Carbon dioxide / oxygen
Environmental conditions	**Indoor fire load** / Circuit aging degree

As aforementioned, the real number attributes and interval attributes can be denoted by TFNs directly. The linguistic attributes "Flammable of storage items" and "Circuit aging degree" can be transformed into TFNs in Table 3.

Table 3: Linguistic attributes and corresponding TFNs

Linguistic terms	TFNs
Not dangerous	(0,0,0.111,0.222)
Medium dangerous	(0.111,0.222,0.333,0.444)
Dangerous	(0.333,0.444,0.555,0.666)
Very dangerous	(0.555,0.666,0.777,0.888)
Highly dangerous	(0.777,0.888,0.999,1)

Then we carry out normalization by Eq. 3 and Eq. 4 to obtain corresponding TFNs.

(2) The first step for screening Scr_1

Without loss of generality, the criterion "Fire behavior" will be utilized to illustrate the first screening procedure and the other two criteria "Personal security" and "Environmental conditions" can be dealt with similarly.

The criterion "Fire behavior" is represented by seven attributes, namely, "Temperature", "Enthalpy", "Heat Release Rate", "Relative Humidity", "Burning Rate", "Pressure" and "Radiative Heat Flux"; and "Temperature" is the "relatively best" representative attribute for the criterion "Fire" behavior.

Step 1.1 to *Step 1.4* are implemented to calculate correlation degrees between ""Temperature" and the remaining six attributes, and the results are given in Table 4.

Table 4: Correlation degrees between "Temperature" and the other six attributes

Attribute	Enthalpy	Heat Release Rate	Relative Humidity	Burning Rate	Pressure	Radiative Heat Flux
Correlation degree	0.53745	0.96987	0.33271	0.74787	0.47665	0.88712

Step 1.5: Set threshold of correlation degree $\tau^* = 0.6$

Thus the attributes will be reserved whose correlation degrees is greater than or equal to 0.6, i.e., "Burning Rate", "Heat Release Rate" and "Radiative Heat Flux".

Step 1.6: Repeat Step 1.1 to Step 1.5 for each criterion

For criteria "Personal security" and "Environmental conditions", the "relatively best" representative attributes are respectively "Carbon monoxide" and "Indoor fire load", and relevant correlation degrees are shown in Table 5.

Table 5: Relevant correlation degrees for the other two criteria

Criterion	Personal security		Environmental conditions
Attribute	Carbon dioxide	oxygen	Circuit aging degree
Correlation degree	0.80961	0.93428	0.76865

We reset thresholds of correlation degree for each criteria, viz., $\tau^*_{Personal} = 0.9$ and $\tau^*_{Environmental} = 0.8$.

Finally, the result of the first screening A_{Scr_1} can be obtained, namely,

A_{Scr_1} = {"Temperature", "Heat Release Rate", "Radiative Heat Flux", "Burning Rate", "Carbon monoxide", "oxygen", "Indoor fire load"}.

(3) The second step for screening Scr_2

Before conducting Scr_2, MU^*, σ and W^* should be given firstly. As aforementioned, MU^* and σ can be determined by relevant research achievements and expert experiences, and therefore, we collected and integrated evaluation results from ten fire researchers to assign MU^* and σ as follows:

$$MU^* = \begin{pmatrix} 19.5 & & & & & & \\ & 15.6 & & & & & \\ & & 7.4 & & & & \\ & & & 11.2 & & & \\ & & & & 15.4 & & \\ & & & & & 18.7 & \\ & & & & & & 18.2 \end{pmatrix} \text{ and } \sigma = \begin{pmatrix} 9 & 8 & 8 & 9 & 4 & 4 & 5 \\ 4 & 5 & 4 & 4 & 9 & 9 & 7 \end{pmatrix}$$

where the elements values of MU^* are in the interval [0, 20], and the greater the value is, the more sensitive the attribute is for evaluation. The elements values of σ are in the interval [0, 10], and the greater the value is, the more familiar the attribute is to DMs.

Moreover, W^* can be acquired with the given criteria weight vector $(\omega_1, \omega_2, \omega_3) = (0.4, 0.4, 0.2)$, namely,

$$W^* = \begin{pmatrix} 0.299 & & & & & & \\ & 0.388 & & & & & \\ & & 0.215 & & & & \\ & & & 0.354 & & & \\ & & & & 0.352 & & \\ & & & & & 0.374 & \\ & & & & & & 0.197 \end{pmatrix}$$

After determine relevant parameters MU^*, W^*, σ, and the seven-dimensional independent variable v^b, Scr_2 based on the multi-objective optimization model Eq 19 is performed in Matlab R2013b. We choose the traditional mathematical programming method to solve this model, because the research [15] has proven that the programming method has better performance than aforementioned other algorithms to solve such a problem with the zero-one binary vector variable and the small size of data.

The results are shown in Figure 5 where the curve represents all possible solutions and the Pareto Optimal solution is the intersection of the curve and the x-axis, namely, $v^b = (1, 1, 0, 1, 1, 1, 1)$. It means that the attribute "Radiative Heat Flux" represented by zero can be omitted, and the optimal set of decision attributes is:

$A_{Scr2} = \{$"Temperature", "Heat Release Rate", "Burning Rate", "Carbon monoxide", "oxygen", "Indoor fire load"$\}$.

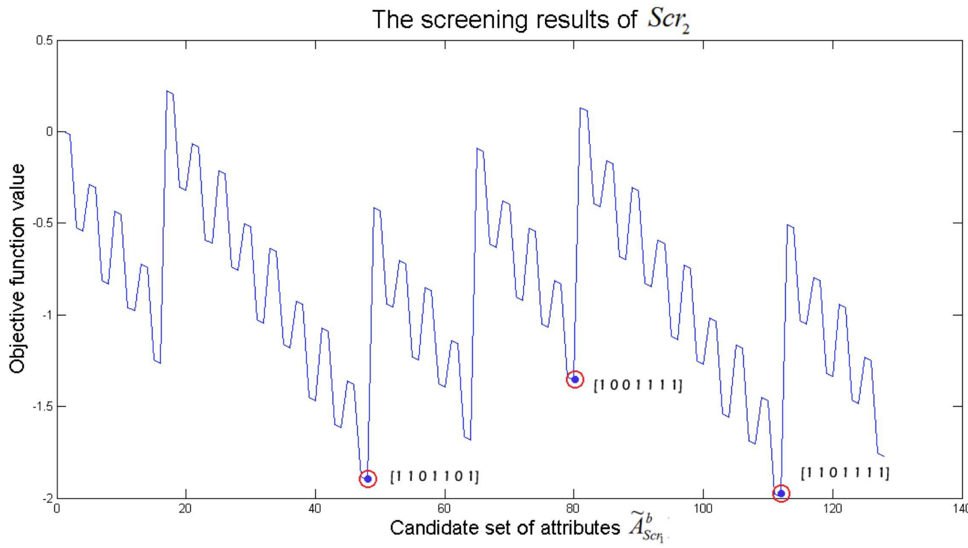

Figure 5: The results of the second step Scr_2

4.3 Validation of the presented method

In order to verify the result is the optimal set for a MCDM problem, multi-attribute utility theory will be used in the following.

A conclusion underlies the presented method: Removing decision-irrelevant attributes from the set of decision attributes will improve the effectiveness of decision alternatives. Thus it is reasonable and feasible to verify our work starting from Scr_2, because the purpose of Scr_1 is to omit decision-irrelevant attributes.

Multi-attribute utility theory (MAUT) is used to choose decision alternatives by evaluating the utility values of alternatives which is calculated based on the utility of decision attributes, and it is an effective way to quantify decision attributes' impact on decision results. More precisely, the utility evaluation model of an alternative is given as follows:

$$u(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} k_i u_i(x_i) \qquad if \sum_{i=1}^{n} k_i = 1 \tag{20}$$

where x_i is the ith decision attribute, and k_i is the corresponding weight; $u_i(x_i)$ is the attribute utility function of x_i, and $u(x_1, x_2, ..., x_n)$ is the utility function of an alternative evaluated by $x_1, x_2, ..., x_n$.

Thus the utility values of all possible candidate sets $\widetilde{A}^b_{Scr_1}$ are calculated based on Eq. 20, where attribute utility function $u_i(x_i)$ can be defined as Eq. 12, and the weights are acquired by Eq. 14. It should be pointed that normalization of weights need be carried out to guarantee $\sum_{i=1}^{n} k_i = 1$. And the results are given in Table 6, where S_i represents the candidate set of decision attributes, and $1 \leq i \leq 16$. The correspondence between S_i and the set of attributes are shown in Table 7.

Table 6: The utility values of attribute sets

Attribute set	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
Utility value	23.621	24.343	28.496	28.923	29.665	31.745	28.474	29.500
Attribute set	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}
Utility value	25.908	28.447	32.631	31.229	30.335	35.774	31.658	33.439

Table 7: The correspondence between S_i and the set of attributes

S_i	The correspondence set	S_i	The correspondence set
S_1	(1, 0, 0, 0, 1, 0, 1)	S_9	(1, 0, 0, 0, 1, 1, 1)
S_2	(1, 1, 0, 0, 1, 0, 1)	S_{10}	(1, 1, 0, 0, 1, 1, 1)
S_3	(1, 0, 1, 0, 1, 0, 1)	S_{11}	(1, 0, 1, 0, 1, 1, 1)
S_4	(1, 0, 0, 1, 1, 0, 1)	S_{12}	(1, 0, 0, 1, 1, 1, 1)
S_5	(1, 1, 1, 0, 1, 0, 1)	S_{13}	(1, 1, 1, 0, 1, 1, 1)
S_6	(1, 1, 0, 1, 1, 0, 1)	S_{14}	(1, 1, 0, 1, 1, 1, 1)
S_7	(1, 0, 1, 1, 1, 0, 1)	S_{15}	(1, 0, 1, 1, 1, 1, 1)
S_8	(1, 1, 1, 1, 1, 0, 1)	S_{16}	(1, 1, 1, 1, 1, 1, 1)

In Table 6, for $\forall i$, $1 \leq i \leq 16$, it has $u(S_{14}) > u(S_i)$, where $i \neq 14$. The set S_{14} has the maximum utility value $u(S_{14}) = 35.774$, and thus the alternative chosen based on S_{14} (the optimal set A_{Scr_2}) is the optimal alternative.

In addition, S_1 contains three attributes with minimum utility value $u(S_1) = 23.621$, and S_{16} contains all attributes with utility value $u(S_{16}) = 33.439$. It means that too few or too many attributes are not available in a MCDM problem; too few attributes are not enough to evaluate the decision problem, while too many attributes may lead to conflict and interfere decision results.

The set of decision attributes obtained by utilizing the presented method has the maximum utility value in alternative evaluation, and it demonstrates the effectiveness of the presented method.

5 Implications

This paper utilizes attribute selection procedure in MCDM problems innovatively. MCDM problems refer to getting the optimal alternatives which are determined based on many qualitative or quantitative criteria, and these criteria are conflicting and assessed by more relevant attributes. Therefore, the set of decision attributes determines the effectiveness of an alternative. In order to improve the reliability of MCDM, the method proposed in this study can select the optimal set of decision attributes which are highly related to the MCDM problem and remove redundant or "noisy" attributes. Meanwhile, the selection of optimal set can reduce the attribute dimension and improve the efficiency of decision making while maintaining the great effect.

The method combines external attribute data with subjective decision preferences in the sequential procedure to improve decision performance. External attribute data are objective factors, which show the features of the data. These attributes contain its own regularity and all kinds of information that decision-making needs, so the first step to screen is based on objective attribute data. Subjective factors show the decision makers' preferences. Because of the different profession and background knowledge, decision makers have different recognition of attributes. Considering the subjective decision preferences contributes to improve the reliability and accuracy of decision-making. Using the former to build the method and using the constraint conditions of the latter to adjust the method can make it more in line with the actual situation.

The method combines external attribute data with subjective decision preferences in the sequential procedure to improve decision performance. External attribute data are objective factors, which show the features of the data. These attributes contain its own regularity and all kinds of information that decision-making needs, so the first step to screen is based on objective attribute data. Subjective factors show the decision makers" preferences. Because of the different profession and background knowledge, decision makers have different recognition of attributes. Considering the subjective decision preferences contributes to improve the reliability and accuracy of decision-making. Using the former to build the method and using the constraint conditions of the latter to adjust the method can make it more in line with the actual situation.

6 Conclusions

In this paper, a new attribute selection method with the context of MCDM is proposed. According to the analysis of attribute selection problem, a two-step procedure is established to reduce original set to the optimal set of decision attributes. GRA theory and multi-objective optimization method are respectively used to implement these two screening steps. And then a fire example is shown to illustrate application and validity of screening method. The attribute selection method presented in this paper is a dynamic and flexible procedure which depends on decision data resource and decision requirement. It eliminates the drawback that decision attributes are chosen completely subjectively, and it will show better performance for new decision scenarios.

Attribute selection is also a critical process for many domains, such as classification problem, clustering problem, risk assessment, credit assessment, etc. And even solution of these issues depended on results of attribute selection. Attribute selection problem is also discussed in data mining called "Feature Engineering", and still has needs to continue thorough research and the technique problem to be solved.

Generally it exists plenty of missing data among decision information, and how to select the optimal decision attributes in this case will be our future research.

Bibliography

[1] Babaei, S., Sepehri, M. M., Babaei, E. (2015). Multi-objective portfolio optimization considering the dependence structure of asset returns. *European Journal of Operational Research*, 244(2), 525-539, 2015.

[2] Bermejo, P., Gamez, J. A., Puerta, J. M. (2014). Speeding up incremental wrapper feature subset selection with Naive Bayes classifier. *Knowledge-Based Systems*, 55, 140-147, 2014.

[3] Chen, Y., Kilgour, D. M., Hipel, K. W. (2008). Screening in multiple criteria decision analysis. *Decision Support Systems*, 45(2), 278-290, 2008.

[4] Chun, Y. H. (2015). Multi-attribute sequential decision problem with optimizing and satisficing attributes. *European Journal of Operational Research*, 243(1), 224-232, 2015.

[5] Comes, T., Hiete, M., Wijngaards, N., Schultmann, F. (2011). Decision maps: A framework for multi-criteria decision support under severe uncertainty. *Decision Support Systems*, 52(1), 108-118, 2011.

[6] Dai, J., Wang, W., Tian, H., Liu, L. (2013). Attribute selection based on a new conditional entropy for incomplete decision systems. *Knowledge-Based Systems*, 39, 207-213, 2013.

[7] Hapfelmeier, A., Ulm, K. (2014). Variable selection by Random Forests using data with missing values, *Computational Statistics* & *Data Analysis*, 80, 129-139, 2014.

[8] Huda, S., Abdollahian, M., Mammadov, M., Yearwood, J., Ahmed, S., Sultan, I. (2014): A hybrid wrapper-filter approach to detect the source (s) of out-of-control signals in multivariate manufacturing process, *European Journal of Operational Research*, 237(3), 857-870, 2014.

[9] Lin, Q., Li, J., Du, Z., Chen, J., Ming, Z. (2015). A novel multi-objective particle swarm optimization with multiple search strategies, *European Journal of Operational Research*, 247(3), 732-744, 2015.

[10] Ma XF, Zhong QY, Qu Y (2013) Determination method of emergency key property based on common knowledge model and Euclidean distance, *Systems Engineering*, 31(10), 93-97, 2013.

[11] Meinshausen, N., Buhlmann, P. (2010): Stability selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(4), 417-473, 2010.

[12] Meng MH, Pei XJ, Wu MQ (2015): Study on choice of factors influencing stability of perilous rock based on fuzzy multi-attribute group decision-making. *Subgrade Engineering*, 1:20-23, 2015.

[13] Montajabiha, M. (2016): An Extended PROMETHE II Multi-Criteria Group Decision Making Technique Based on Intuitionistic Fuzzy Logic for Sustainable Energy Planning. *Group Decision and Negotiation*, 25(2), 221-244, 2016.

[14] Robin, G., Jean-Michel P., Christine T. (2010): Variable selection using random forests, *Pattern Recognition Letters*, 31, 2225-2236, 2010.

[15] Shen HP, Zhang YP, Wang YK (2014): Research on regular Chinese fragments reassembly based on 0-1 programming model, *Electronic Science and Technology*, 6:13-16, 2014.

[16] Stewart, T. J. (1992): A critical survey on the status of multiple criteria decision making theory and practice, *Omega*, 20(5-6), 569-586, 1992.

[17] Wahlqvist, J., Van Hees, P. (2013): Validation of FDS for large-scale well-confined mechanically ventilated fire scenarios with emphasis on predicting ventilation system behavior, *Fire Safety Journal*, 62, 102-114, 2013.

[18] Wu, K. J., Tseng, M. L., Chiu, A. S., Lim, M. K. (2016): Achieving competitive advantage through supply chain agility under uncertainty: A novel multi-criteria decision-making structure, *International Journal of Production Economics*, article in press, 2016.

[19] Wu, K. J., Liao, C. J., Tseng, M. L., Lim, M. K., Hu, J., Tan, K. (2017): Toward sustainability: using big data to explore the decisive attributes of supply chain risks and uncertainties, *Journal of Cleaner Production*, 142, 663-676, 2017.

[20] Zeleny, M., Cochrane, J. L. (1973): *Multiple criteria decision making*, University of South Carolina Press, 1973.

[21] Zhang, Y., Gong, D., Cheng, J. (2017): Multi-Objective Particle Swarm Optimization Approach for Cost-based Feature Selection in Classification, *IEEe/ACM Transactions on Computational Biology and Bioinformatics*, 14, 64-75, 2017.

[22] Zhu, J., Zhang, S., Chen, Y., Zhang, L. (2016): A hierarchical clustering approach based on three- dimensional gray relational analysis for clustering a large group of decision makers with double information, *Group Decision and Negotiation*, 25(2), 325-354, 2016.