Impute Missing Assessments by Opinion Clustering in Multi-Criteria Group Decision Making Problems

Jun Ma[1], Guangquan Zhang[1], Jie Lu[1] Da Ruan[2,3]

1. School of Software Engineering, Faculty of Engineering and Information Technology
University of Technology, Sydney (UTS), PO Box 123, Broadway, NSW 2007, Australia
2. Beglian Nuclear Research Centre (SCK•CEN), 2400, Boeretong, Mol, Belgium
3. Department of Applied Math. & Computer Science, Ghent University, 9000 Gent, Belgium
Email: {junm, zhangg, jielu}@it.uts.edu.au, druan@sckcen.be, da.ruan@ugent.be

Abstract— Multi-criteria group decision-making and evaluation (MCGDME) method typically aggregates information in evaluation tables. For various reasons, evaluation tables (decision matrix) often include missing data that highly affect correct decision-making and evaluation. Most existing imputation methods of missing data are based on statistical features which do not exist in an MCGDME setting. This paper proposes an imputation method of missing data (IMD) in evaluation tables. The IMD method measures the similarity betweent two evaluators' mental models. Evaluators are then classed into several groups based on their similarities by using fuzzy clustering methods. Finally, missing data are imputated under the assumption that the imputated value of missing data does not change the previous clustering results. The proposed IMD method is implemented and tested in two numerical experiments.

Keywords— decision making, missing data, multi-criteria evaluation, opinion clustering, aggregation

1 Introduction

Multi-criteria group decision-making (MCGDM) methods have been widely used in various fields such as new product development in industries [9], clinical diagnose, service strategies selection in business, as well as public economic policy evaluation [11]. In a typical MCGDM process, evaluating alternatives and aggregating assessments based on decision matrices are two crucial processes. Evaluators (or experts) assess alternatives in terms of a set of identified criteria. The assessments are recorded into several decision matrices, also known as decision tables, which are further analyzed and integrated. The aggregation process implements a specific aggregation model to achieve an overall assessment for each alternative and based on which the best alternative(s) is/are selected. Majority of existing research focuses particularly on the problem of how to efficiently integrate heterogeneous quantitative or qualitative information from multiple information sources. Most processes are conducted mainly on the assumption that those obtained decision matrices are intact, i.e., those tables do not include missing values. However, this assumption may not be completely satisfied in real situations.

Various reasons result in missing values in decision matrices. Values may be missed in the course of data collection by device faults or human errors. Evaluators who are lack of relevant background for some criteria or are not willing to provide opinions on some criteria may leave blanks in their decision tables. Security or privacy policies may also prevent evaluators from providing answers to some criteria. An evidently result of missing values reduce the reliability of decision making or evaluation and to produce extraordinary bias towards appropriate decisions. Moreover, missing values may result in invalid evaluation and decision making. Hence, necessary and efficient missing data imputation is required in real applications such as in nuclear material safeguards information management.

Literature indicates missing data processing has a long research history [4–6, 8, 12]. However, most research is mainly conducted on the theory and application of statistics sciences [5]. Existing techniques, such as the LOCF (last observation carried forward), best or worst case imputation, mean imputation, and case deletion [1], generally require the collected data has specific and identifiable features. Such features include, for examples, the data is time-related or follows a particular statistics distribution, as well as has sufficient volume. These features do not hold in a typical MCGDM settings in general. Assessments in a typical MCGDM problem have uniqueness in the sense that each assessment is uniquely linked to a unique alternative, a unique evaluator, and a unique criterion. Moreover, because evaluators are generally not asked or are not willing to repeat same evaluation procedures, the obtained assessment values are always onetime data which is irrelevant to observation time. Therefore, not enough volume of data exists with the same features of a missing value. Besides the uniqueness and small volume of assessments, assessment values are generated through information processing in evaluators' mental model of the given decision problem which cannot be strictly expressed through a formalized and quantified model. Moreover, the links among the selected criteria are generally weak and unclear, because evaluation criteria are always deliberately selected to represent different considerations for alternatives. These features make it very hard to use statistics-based missing data imputation methods to MCGDM problems.

This paper presents an imputation method for missing assessments (IMD) in decision tables for MCGDM. The presented method mainly includes two steps. A fuzzy clustering method is first used to classify evaluators into several groups based on their opinions reflected in their assessments. Then missing assessments are imputed according to those assessments of evaluators belonging to the same group. The rest of the paper is organized as follows. Section 2 reviews the main steps of the presented IMD method. Details of each step is introduced in Section 3. To further illustrate the IMD method,

ISBN: 978-989-95079-6-8

two case studies are conducted. Finally, main results and future works are discussed in Section 5.

2 Overview

2.1 Formalization of typical MCGDM problems

A typical MCGDM problem is composed of a set of alternatives (e.g., action choices, policies), a hierarchy of attributes (i.e., decision criteria), weights associated with criteria, a set of decision matrices (i.e., decision tables), as well as a set of evaluators who present those decision matrices. Three main steps are involved in selections of alternatives [13]: 1) determine the relevant criteria and alternatives; 2) evaluate the relative impacts of alternatives on those criteria; and 3) determine a ranking of each alternatives. Roughly, an MCGDM process can be expressed by a 4-tuple model:

$$\mathcal{M} = (\mathcal{C}, \mathcal{E}, \mathcal{A}, \mathcal{T}), \tag{1}$$

where $\mathcal{C} = \{(c_j, wc_j)|j = 1, 2, \ldots, n\}$ is a set of criteria and their corresponding weights (importance); $\mathcal{E} = \{(e_k, we_k)|k = 1, 2, \ldots, m\}$ is a set of evaluators (experts) and their corresponding weights (reliabilities); $\mathcal{A} = \{a_i|i = 1, 2, \ldots, p\}$ is a set of alternatives (action choices); and $\mathcal{T} = \{T_i = (v^i_{jk})_{n\times m}|i = 1, 2, \ldots, p\}$ is a set of decision matrices. For each a_i, the overall assessment y_i based on the model $\mathcal{M}$ is explicitly expressed as

$$y_i = W_C \circ T_i \diamond W_E^{\mathrm{T}} = (wc_1, \cdots, wc_n) \circ \begin{pmatrix} v^i_{11} & v^i_{12} & \cdots & v^i_{1m} \\ v^i_{21} & v^i_{22} & \cdots & v^i_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ v^i_{n1} & v^i_{n2} & \cdots & v^i_{nm} \end{pmatrix} \diamond \begin{pmatrix} we_1 \\ we_2 \\ \vdots \\ we_m \end{pmatrix}$$

where $\circ$ and $\diamond$ are operations in $\mathcal{M}$, W_C is the weight vector for criteria, and W_E the weight vector for evaluators. Once $\circ$ and $\diamond$ are given in explicit forms, the model $\mathcal{M}$ is then determined. For example, let $\circ$ and $\diamond$ be a weighted sum. For any vectors $\boldsymbol{w}$ and $\boldsymbol{v}$

$$\boldsymbol{w} \circ \boldsymbol{v}^{\mathrm{T}} = (w_1, \cdots, w_n) \circ (v_1, \cdots, v_n)^{\mathrm{T}} = \sum_{i=1}^{n} w_i v_i, \tag{2}$$

then y_i can be written as

$$y_i = \sum_{j=1}^{n} \left(wc_i \cdot \sum_{k=1}^{m} v_{jk} we_k \right) = \sum_{k=1}^{m} \left(\sum_{j=1}^{n} wc_j v_{jk} \right) \cdot we_k. \tag{3}$$

Note that most research works on MCGDM focus on how to define these two operations which are mostly studied in the framework of aggregation operators.

In the following, let $T_i = (\boldsymbol{v}^i_1, \boldsymbol{v}^i_2, \cdots, \boldsymbol{v}^i_k)$, where $\boldsymbol{v}^i_k = (v^i_{1k}, v^i_{2k}, \cdots, v^i_{nk})^{\mathrm{T}}$ is called the evaluation vector of evaluator e_k on alternative a_i. T_i is called complete if no missing values in T_i. T_i is called incomplete if at least one evaluation vector includes missing values. For convenience, the following analysis will omit the alternative index i.

2.2 Distance and similarity of evaluator opinions

For each $\boldsymbol{v}_k$ in T, $\boldsymbol{v}_k$ can be seen as an output of evaluator e_k's mental model with the physical nature of alternative as input and denoted by x, i.e., this can be formally described as

$$\boldsymbol{v}_k = f_k(x), \quad k = 1, 2, \ldots, m. \tag{4}$$

Figure 1 illustrates a generalized procedure of obtaining an evaluation table. Notice that the input of each evaluator's mental model is the same, the difference among the obtained evaluation vectors can only be introduced by the difference among evaluators' mental models. Assume that the relationship between evaluators' mental models exists and can be found through their outputs, i.e., the evaluation vectors, then missing values can be approximately imputed with the aid of those relationships. In the light of this idea, the following strategy is adopted to impute missing values. First, a distance measure d between two evaluation vectors is defined, which is then taken as the distance between two evaluators' opinions. Next, the similarity of two evaluators' mental models is induced from the distance. Based on the similarities between each pair of evaluators, a fuzzy clustering method is used to segment evaluators to several clusters. For each of missing values, a predicted value is obtained according to the evaluation vectors provided by evaluators belonging to the same group. This value is the imputed value.

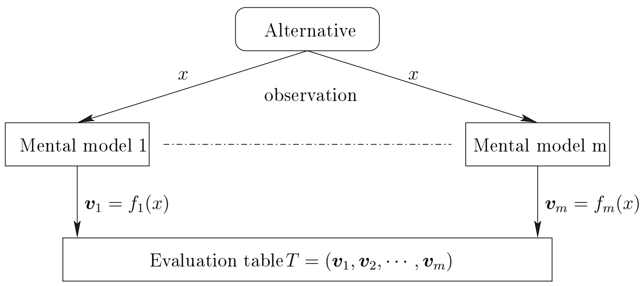

Figure 1: Producing of an evaluation table.

To implement the above-mentioned strategy, we consider some relevant issues. First, the difference of criteria weights is considered when defining a distance measurement. The weight of a criterion indicates the importance degree or relevance degree of that criterion to the decision goal. Some criteria are more important than others. Hence, criteria with bigger weights should contribute more distance to the overall distance between two evaluation vectors. For this reason, we use an aggregation operator to define the overall distance between two evaluation vectors. Another issue is concerned about criteria to that not all missing values must be imputed. For criteria with very small weights, the imputation of missing data is not essential because the bias from appropriate decision induced by discrepancy in those criteria is insignificant. Thus, a parameter α is defined to filter criteria with small weights. Missing data with respect to those filtered criteria are not imputed. Moreover, a second parameter β is used in the segmentation of evaluators, which indicates the at least similarity required when two evaluators are treated with the similar opinion. It can be proved that the bigger the β is, the more obtained clusters and the reduced number of evaluators belong to a certain cluster.

Because the presented strategy adopts two extra parameters α and β, the presented method is thus called an α-β imputation. Next section will give details of the α-β imputation method.

ISBN: 978-989-95079-6-8

3 An α-β imputation method

This section illustrates the details of the α-β imputation method for missing data (assessments, evaluations). For convenience, let all weights for criteria and evaluators be real numbers in the real unit interval [0, 1] and all assessment values be real numbers in [0, 1], too. Without loss of generality, the following discussion will consider the decision matrix for only one alternative.

3.1 Distance between evaluation vectors

An evaluation vector refers to a column in a decision matrix T^i which records the assessments presented by a certain evaluator. As each evaluation vector reflects the opinion of an evaluator, distance of evaluation vectors could be used to reflect the difference between two evaluators' opinions. Based on this idea, we first measure the distance between evaluation vectors. Suppose $\boldsymbol{u}$ and $\boldsymbol{v}$ are two evaluation vectors.

For each criterion c_j ($j = 1, 2, \ldots, n$), let u and v be two assessments from different evaluators about c_j. Then the distance between u and v can be calculated. Suppose the distance is calculated through a distance function d_j which is with respect to c_j and $d_j(u, v)$ is a real number in $[0, 1]$, then d_j can be treated as a function from $[0, 1] \times [0, 1]$ to $[0, 1]$. In the following, let $d_1, d_2, \ldots, d_n$ be the selected distance functions with respect to criteria $c_1, c_2, \ldots, c_n$, respectively without other specifications. Here, we do not require that all d_j be of the same form.

Based on the selected d_j, $j = 1, 2, \ldots, n$, the overall distance between two evaluation vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ can be calculated as follows. Let Agg be an aggregation operator [10] defined from $[0, 1]^n$ to $[0, 1]$ and $\beta_1, \beta_2, \ldots, \beta_n$ be the normalized weights of those criteria, i.e.,

$$\beta_j = \frac{wc_j}{\sum_{j=1}^{n} wc_j}, \quad j = 1, 2, \ldots, n. \tag{5}$$

The distance function $\hat{d}$ between the two evaluation vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ is a function from $[0, 1]^n$ to $[0, 1]$ such that

$$\hat{d}(\boldsymbol{u}, \boldsymbol{v}) = \text{Agg}\left\{(\beta_j, d_j(u_j, v_j)) | j = 1, 2, \ldots, n\right\}. \tag{6}$$

Example 3.1 *Suppose Table 1 gives an example setting. Let* Agg *be the weighted-mean. Then the distance between evaluation vectors* $\boldsymbol{u}$ *and* $\boldsymbol{v}$ *is:*

$$\frac{0.12}{1.52} \times 0.75 + \frac{0.47}{1.52} \times 0.11 + \frac{0.93}{1.52} \times 0.57 = 0.44.$$

If Agg *is the OWA operator, then the distance becomes*

$$\frac{0.12}{1.52} \times 0.75 + \frac{0.47}{1.52} \times 0.57 + \frac{0.93}{1.52} \times 0.11 = 0.30.$$

Table 1: An example settings

	d_j	wc_j	$\boldsymbol{u}$	$\boldsymbol{v}$
c_1	$\lvert u - v\rvert$	0.12	0.07	0.82
c_2	$\lvert u - v\rvert^2$	0.47	0.64	0.97
c_3	$\lvert u - v\rvert^3$	0.93	0.93	0.1

Equation (6) indicates the distance between two evaluation vectors that are determined by the selected aggregation operator and the distance functions for each of criteria. It is possible to choose different operators and measures according to the real problem.

3.2 Similarity between two evaluation vectors

Similar to the definition of distance, similarity for individual criterion is first given and then an overall similarity is calculated by a selected aggregation operator.

Let g_j ($j = 1, 2, \ldots, n$) be the selected similarity function with respect to criterion c_j. g_j should satisfy the following two conditions: for any u and v,

$$g_{j_1}(u, v) \geqslant g_{j_2}(u, v) \text{ if } wc_{j_1} \leqslant wc_{j_2}, \quad j_1, j_2 \in \{1, \cdots, n\}; \tag{7}$$

and g_j is a non-increasing function with respect to d_j. Then function g_j is also called a similarity measure with respect to criterion c_j, $j = 1, 2, \ldots, n$. (7) indicates that a more important criterion has less similarity with respect to the same distance between two assessments.

For example, the following functions can be treated as a set of similarity measures.

$$g_j(u, v) = 1 - d_j(u, v)^{\sigma / wc_j}, \quad j = 1, 2, \ldots, n, \tag{8}$$

where $\sigma \in \{wc_j | j = 1, 2, \ldots, n\}$.

In the light of individual similarity measures $\{g_j\}$, an overall similarity $\hat{g}$ between two evaluation vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ is defined, which is a non-increasing function from $[0, 1]^n$ to $[0, 1]$ such that

$$\hat{g}(\boldsymbol{u}, \boldsymbol{v}) = \text{Agg}\{(\beta_j, g_j(u_j, v_j)), j = 1, 2, \ldots, n\}, \tag{9}$$

where Agg is a selected aggregation operators, β_j is the normalized weights of those criteria, $j = 1, 2, \ldots, n$.

In Example 3.1, suppose $g_j(u, v) = 1 - d_j(u, v)$, then the similarities for the weighted-sum aggregation and OWA aggregation are 0.56 and 0.70, respectively. Obviously $\hat{g}(\boldsymbol{u}, \boldsymbol{v}) = 1 - \hat{d}(\boldsymbol{u}, \boldsymbol{v})$. However, this does not hold if we change the relationship between g_j and d_j for any j. This indicates that in real problems we can select the appropriate operators to define the similarity. Moreover, we should notice that $\hat{g}(\boldsymbol{u}, \boldsymbol{v})$ is given through individual distance d_j and individual similarity g_j, $j = 1, 2, \ldots, n$, rather than the overall distance $\hat{d}(\boldsymbol{u}, \boldsymbol{v})$. The reason for separately defining them is to provide more flexibilities to consider the individual influence of criteria.

3.3 Evaluator segmentation

The similarity between two evaluation vectors reflects to what extend two evaluators' opinions are similar, which is used to define the similarity between two evaluators' mental models. Evaluators, thus, can be clustered into several groups based on the similarity of their mental models.

Suppose the similarities between any two pair of evaluators are calculated and recorded in a similarity matrix $M = (s_{jk})_{m \times m}$ where s_{jk} is the similarity between evaluators e_j and e_k. To segment evaluators based on M, we use the fuzzy clustering technique here. Let $\bar{M}$ be the transitive closure of M obtained by

$$\bar{M} = \bigcup_{i=1} M^i, \tag{10}$$

where M^i is the i-th power of M by the max-min composition [11]. Therefore, a clustering result can be obtained from the

ISBN: 978-989-95079-6-8

matrix $\bar{M}$. For example, suppose the matrix $\bar{M}$ is

$$\bar{M} = \begin{pmatrix} 1.0 & 0.6 & 0.8 & 0.6 \\ 0.6 & 1.0 & 0.6 & 0.7 \\ 0.8 & 0.6 & 1.0 & 0.6 \\ 0.6 & 0.7 & 0.6 & 1.0 \end{pmatrix}. \quad (11)$$

Then the clustering result is shown in the following figure.

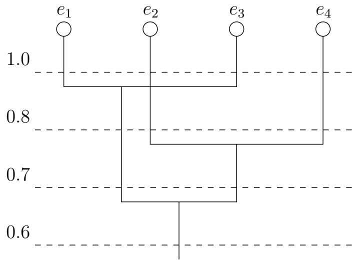

Figure 2: An evaluator clustering results.

From Figure 2, the segmentation results may change with different levels of similarity. For instance, the segmentation result at $\hat{g} = 0.8$ is $\{e_1, e_3\}$, $\{e_2\}$, and $\{e_4\}$; while all evaluators belong to the unique group at $\hat{g} = 0.6$. Hence, a parameter β is used here to indicate at which level the segmentation result is concerned about. In the following let the segmentation result with respect to the given β be $\Gamma(\beta)$.

3.4 *Imputation of missing data*

To impute missing data, we consider relevant issues and illustrate corresponding solutions here.

First, some missing data need not to be imputed when the associated criteria have infinitesimal effects to the decision G. Therefore, a parameter α is introduced to filter criteria with smaller weights than α. For those criteria, missing data are not imputed.

Next, the evaluator segmentation in Section 3.3 is conducted under the assumption that the decision matrix is a complete one. When the decision matrix is incomplete, we make the following adjustment. Without loss of generality, set

$$T = \begin{pmatrix} A \\ B \end{pmatrix}, \quad (12)$$

where A is a sub-matrix of T which is complete and B is the remainder of T in which each row contains at least one missing data. The evaluator segmentation is obtained based on A, i.e., taking A as the basis to calculate distance and similarity between evaluation vectors. Values in B are used to generate imputations of missing values. Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be two evaluation vectors and divide them into two parts according to A, B,

$$\boldsymbol{u} = \begin{pmatrix} u_A \\ u_B \end{pmatrix}, \qquad \boldsymbol{v} = \begin{pmatrix} v_A \\ v_B \end{pmatrix}.$$

Suppose $u_j \in u_B$ is a missing value and $v_j \in v_B$ exists. To estimate u_j from v_j, we follow the principle that adding missing value does not change the established similarity between the two evaluation vectors. That means if the established similarity between u_A and v_A is g, then the similarity between $u_{A\cup\{u_j\}}$ and $v_{A\cup\{v_j\}}$ is still g after imputation the missing value u_j. Formally, this can be expressed as the following equation

$$\begin{aligned} &\text{Agg}\Big\{(\beta'_l, g_l(u_l, v_l)), l \in \delta(A) \cup \{j\}\Big\} \\ &\quad = \text{Agg}\Big\{(\beta_l, g_l(u_l, v_l)), l \in \delta(A)\Big\}, \end{aligned} \quad (13)$$

where $\delta(A) = \{l | u_l \in A\}$, β' is the normalized weight after adding the criterion c_j and is calculated by

$$\beta'_l = \frac{wc_l}{\sum_{r\in\delta(A)\cup\{j\}} wc_r} = \frac{wc_l}{wc_j + \sum_{r\in\delta(A)} wc_r}. \quad (14)$$

Notice in (13) only u_j is unknown, therefore, the solution(s) of the equation is/are possible estimation(s) of the missing value u_j in $\boldsymbol{u}$ in terms of $\boldsymbol{v}$.

For example, suppose the Agg is the weighted mean and $g_j(\boldsymbol{u}, \boldsymbol{v}) = 1 - d_j(\boldsymbol{u}, \boldsymbol{v})$, $j = 1, 2, \ldots, n$. Then by (13), we have

$$\begin{aligned} &\text{Agg}\Big\{(\beta'_l, g_l(u_l, v_l)), l \in \delta(A) \cup \{j\}\Big\} \\ &\quad = \text{Agg}\Big\{(\beta_l, g_l(u_l, v_l)), l \in \delta(A)\Big\}, \end{aligned}$$

Hence,

$$\sum_{l\in\delta A\cup\{j\}} \beta'_l g_l(u_l, v_l) = \sum_{l\in\delta(A)} \beta_l g_l(u_l, v_l),$$

$$\sum_{l\in\delta(A)\cup\{j\}} \frac{wc_l \cdot g_l(u_l, v_l)}{wc_j + \sum_{r\in\delta(A)} wc_r} = \sum_{l\in\delta(A)} \frac{wc_l \cdot g_l(u_l, v_l)}{\sum_{r\in\delta(A)} wc_r},$$

$$1 - \sum_{l\in\delta(A)\cup\{j\}} \frac{wc_l \cdot d_l(u_l, v_l)}{wc_j + \sum_{r\in\delta(A)} wc_r} = 1 - \sum_{l\in\delta(A)} \frac{wc_l \cdot d_l(u_l, v_l)}{\sum_{r\in\delta(A)} wc_r}$$

So,

$$d_j(u_j, v_j) = \frac{\sum_{l\in\delta(A)} wc_l \cdot d_l(u_l, v_l)}{\sum_{r\in\delta(A)} wc_r} = \hat{d}(u_A, v_A). \quad (15)$$

Equation (15) indicates that u_j can be obtained simply by the d_j and $\hat{d}$ under the above assumptions. Notice that (15) is obtained from g_j and $\hat{g}$, this equation also indicates that the estimation of u_j is determined by the selection of g_j and $\hat{g}$. Hence given different definitions of them, the relationship in (15) may vary.

3.5 *Discussions*

To find solution of (15), we notice that the equation may produce more than one estimation of the missing value. The question, then, raises naturally that how to determine a unique one. This paper partly resolves this problem through calculating the change $\tilde{d}$ of overall distance induced by an obtained estimation. The estimation that produces the least $\tilde{d}$ will be taken as the imputation of missing data. The overall distance change $\tilde{d}$ is defined as

$$\tilde{d}(u_j) = \sum_{v\in\Gamma(u)\setminus\{u\}} d_j(u_j, v_j), \quad (16)$$

where $\Gamma(u)$ is the cluster of evaluation vectors in which u belongs to. By (16), the best estimation is the one which produces the least distance sum between itself to other evaluation

ISBN: 978-989-95079-6-8

vectors belonging to the same cluster. Examples in the following section have shown that this method can reduce some estimations although they cannot grantee the uniqueness of estimation. Hence, more work is being done to provide other strategy to resolve this problem.

Another question raising from the above imputation procedure is the obtained evaluator segmentation may not be correct because the segmentation is conducted based on A only. Accounting for this problem, we extend the above procedure as follows. For any $\boldsymbol{u}$ and $\boldsymbol{v}$, the distance $\hat{d}(\boldsymbol{u}, \boldsymbol{v})$ and the similarity $\hat{g}(\boldsymbol{u}, \boldsymbol{v})$ are calculated in terms of criteria on which both $\boldsymbol{u}$ and $\boldsymbol{v}$ are not missed rather than limited to criteria corresponding to A. That means the obtained distance and similarity between $\boldsymbol{u}$ and $\boldsymbol{v}$ are the most possible values from the current incomplete decision table. An experiment of the extended process is given in the next section.

4 Experiments and observations

The presented α-β imputation method of missing data is programmed in the Java^{TM} programming language running on a Linux system.

4.1 Experiment 1

Experiment 1 is conducted on a complete decision table with 16 criteria and 5 evaluators. Missing data are randomly selected from the decision table.

To impute the missing data, we set the distance measure for each criteria as $d_j(u_j, v_j) = |u_j - v_j|$ $(j = 1, 2, \ldots, 16)$ and the similarity measure for each criterion as $g_j(u_j, v_j) = 1 - d_j(u_j, v_j)$, the aggregation operators used in calculating overall distance and similarity are the same, say the weighted mean; and set the two parameters $\alpha = 0.2$ and $\beta = 0.6$.

Table 2: The original evaluation table in which missing data are marked by bold font.

wc_j	e_1	e_2	e_3	e_4	e_5
0.262	0.877	0.035	0.88	0.772	0.941
0.334	0.977	0.458	0.903	0.898	0.227
0.823	0.845	0.065	0.894	0.122	0.555
0.216	0.532	0.630	**0.676**	0.353	0.858
0.803	0.862	0.073	0.183	0.618	0.784
0.227	0.797	0.046	**0.721**	0.551	0.950
0.794	0.159	0.995	0.621	0.808	0.307
0.811	0.227	0.936	0.994	0.973	0.134
0.945	0.304	0.901	0.636	0.225	**0.927**
0.892	0.967	0.717	0.851	0.835	**0.035**
0.344	0.641	**0.336**	0.666	0.281	0.598
0.016	**0.475**	0.915	0.743	0.876	0.651
0.235	0.026	0.999	0.501	0.431	0.586
0.905	0.307	0.371	0.757	0.832	0.66
0.316	0.541	0.003	0.438	0.113	0.275
0.068	0.042	0.27	0.929	0.675	0.206

By the given distance and similarity measure, we get the distance matrix $\mathcal{D}$ and the similarity matrix $\mathcal{M}$ as follows:

$$\mathcal{D} = \begin{pmatrix} 0 & 0.571 & 0.35 & 0.392 & 0.233 \\ 0.571 & 0 & 0.348 & 0.332 & 0.55 \\ 0.35 & 0.348 & 0 & 0.255 & 0.382 \\ 0.392 & 0.332 & 0.255 & 0 & 0.394 \\ 0.233 & 0.55 & 0.382 & 0.394 & 0 \end{pmatrix} \quad (17)$$

and

$$\mathcal{M} = \begin{pmatrix} 1 & 0.429 & 0.65 & 0.608 & 0.767 \\ 0.429 & 1 & 0.652 & 0.668 & 0.45 \\ 0.65 & 0.652 & 1 & 0.745 & 0.618 \\ 0.608 & 0.668 & 0.745 & 1 & 0.606 \\ 0.767 & 0.45 & 0.618 & 0.606 & 1 \end{pmatrix}. \quad (18)$$

By the fuzzy clustering method, we get the following clustering result.

$$\bar{\mathcal{M}} = \begin{pmatrix} 1 & 0.65 & 0.65 & 0.65 & 0.767 \\ 0.65 & 1 & 0.668 & 0.668 & 0.65 \\ 0.65 & 0.668 & 1 & 0.745 & 0.65 \\ 0.65 & 0.668 & 0.745 & 1 & 0.65 \\ 0.767 & 0.65 & 0.65 & 0.65 & 1 \end{pmatrix} \quad (19)$$

i.e., we have only one class $\{e_1, e_2, e_3, e_4, e_5\}$. By the imputation method in Section 3.4, the following result is obtained (shown in Table 3).

Table 3: Decision table after imputation of missing data (bold number for estimations and # for still missing data).

wc_j	e_1	e_2	e_3	e_4	e_5
0.262	0.877	0.035	0.88	0.772	0.941
0.334	0.977	0.458	0.903	0.898	0.227
0.823	0.845	0.065	0.894	0.122	0.555
0.216	0.532	0.63	**0.608**	0.353	0.858
0.803	0.862	0.073	0.183	0.618	0.784
0.227	0.797	0.046	**0.296**	0.551	0.95
0.794	0.159	0.995	0.621	0.808	0.307
0.811	0.227	0.936	0.994	0.973	0.134
0.945	0.304	0.901	0.636	0.225	**0.071**
0.892	0.967	0.717	0.851	0.835	**0.734**
0.344	0.641	#	0.666	0.281	0.598
0.016	**0.418**	0.915	0.743	0.876	0.651
0.235	0.026	0.999	0.501	0.431	0.586
0.905	0.307	0.371	0.757	0.832	0.66
0.316	0.541	0.003	0.438	0.113	0.275
0.068	0.042	0.27	0.929	0.675	0.206

4.2 Experiment 2

In the presented imputation method of missing data, aggregation operators are used for calculation of overall distance and overall similarity of two evaluation vectors. Different aggregation operators may lead to different clustering results and, in turn, may result in different imputations. Experiment 2 compares three widely-used aggregation operators, i.e., the weighted mean (Agg_w); the arithmetic mean (Agg_a)

$$\mathrm{Agg}_a(x_1, x_2, \cdots, x_n) = \frac{1}{n} \sum_{i=1}^{n} x_i; \quad (20)$$

and the OWA operator (Agg_o)

$$\mathrm{Agg}_o(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} w_i x_{\sigma(j)}, \quad (21)$$

where σ is a permutation such that $x_{\sigma(1)} \geqslant x_{\sigma(2)} \geqslant \cdots \geqslant x_{\sigma n}$. To conduct this experiment, we change the distance measure and the similarity measure for each individual criterion,

ISBN: 978-989-95079-6-8

i.e.,

$$d_j(u_j, v_j) = |u_j - v_j|, \quad g_j(u_j, v_j) = 1 - d_j(u_j, v_j).$$

The original evaluation table is still the one shown in Table 2. The comparison results are shown in Table 4.

Table 4: Comparison results.

missing data		original value	imputations		
row	col		Agg_w	Agg_o	Agg_a
3	2	0.676	0.608	0.527	0.581
5	2	0.721	0.296	0.756	0.323
8	4	0.927	0.071	/	0.557
9	4	0.035	0.734	/	0.714
10	1	0.336	/	/	/
11	0	0.475	0.418	/	0.398

From the comparison, we can observe that different aggregation operators may produce different imputations. Even more, some aggregation operators may not produce estimation of missing data. Moreover, the bias between an imputation and an original value is bigger for some criteria. The independence nature of the data and the unsupervised nature of the presented method are two possible reasons for these results.

5 Conclusions and Future Works

Imputation of missing data is a vital step in decision making and evaluation problems. Due to various reasons, decision tables are often including missing data. Therefore, how to efficient impute the missed data will affect correct decision making and evaluation. However, few work is reported on imputation of missing data for decision tables. In this paper, an α-β imputation method of missing data (IMD) in a decision table is proposed. The IMD method estimates missing data based on the similarity between evaluators' mental models. First, the IMD method uses an overall distance to measure the similarity of two evaluators mental models through their evaluation vectors. Then a fuzzy clustering method is used to segment evaluators based on the similarity of their mental models. Finally, missing data are imputed by the clustering results under the assumption that an imputed data should not change the clustering results. Two numerical experiments are conducted to test the proposed IMD method. Experiments indicate the IMD method can efficiently estimate some missing data. Moreover, those experiments also show that selecting different aggregation operators, distance measure, and similarity measure may change the imputed values. However, from these experiments, we notice that some estimations of missing data have bigger bias from the originals due to the independence nature of the data and lack of backgrounds as well as other possible reasons. Hence, more work still need to be done such as how to improve the accuracy of an estimation of missing data and how to combine relevant background to the imputation. Now, some related work is undertaking.

Currently, we notice that another missing value processing problem in fuzzy preference relations for group decision-making problem has been discussed in [2, 3]. Method based on additive consistency has been discussed [7]. Although the missing value in that method is different from the one discussed in the presented work, ideas in those work can be introduced into the presented work to improve the efficiency and accuracy. Some work is conducting.

Acknowledgment

The work presented in this paper was supported by Australian Research Council (ARC) under Discovery Projects DP0559213 and DP0880739.

The authors sincerely appreciate the reviewers comments and suggestions to improve the presented work.

References

[1] Points to consider on missing data, 2001. Committee for Proprietary Medicinal Products (CPMP),The European Agency for the Evaluation of Medicinal Products.

[2] S. Alonso, F. J. Cabrerizo, F. Chiclana, F. Herrera, and E. Herrera-Viedma. Group decision making with incomplete fuzzy linguistic preference relations. *International Journal of Intelligent Systems*, 24(2):201–222, 2009.

[3] S. Alonso, F. Herrera-Viedma, F. Chiclana, and F. Herrera. Individual and social strategies to deal with ignorance situations in multi-person decision making. *International Journal of Information Technology and Decision Making*, 2009. in press.

[4] Q. Chen, J. G. Ibrahim, M.-H. Chen, and P. Senchaudhuri. Theory and inference for regression models with missing responses and covariates. *Journal of Multivariate Analysis*, 99:1302–1331, 2008.

[5] A. Farhangfar, L. Kurgan, and J. Dy. Impact of imputation of missing values on classification error for discrete data. *Pattern Recognition*, 41:3692–3705, 2008.

[6] O. Harel. Inferences on missing information under multiple imputation and two-stage multiple imputation. *Statistical Methodology*, 4(1):75–89, January 2007.

[7] E. Herrera-Viedma, F. Chiclana, F. Herrera, and S. Alonso. Group decision-making model with incomplete fuzzy preference relations based on additive consistency. *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, 37(1):176–189, 2007.

[8] J. Liesiö, P. Mild, and A. Salo. Robust portfolio modeling with incomplete cost information and project interdependencies. *European Journal of Operational Research*, 190:679–695, 2008.

[9] J. Lu, G. Zhang, D. Ruan, and F. Wu. *Multi-Objective Group Decision Making – Methods, Software and Applications with Fuzzy Set Technology*. Imperial College Press, London, 2007.

[10] R. Mesiar, A. Kolesárová, T. Calvo, and M. Komorníková. A review of aggregation functions. In H. Bustince, F. Herrera, and J. Montero, editors, *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models*, volume 220 of *Studies in Fuzziness and Soft Computing*, pages 121–144. Springer Berlin/Heidelberg, 2008.

[11] G. Munda. A conflict analysis approach for illuminating distributional issues in sustainability policy. *European Journal of Operational Research*, 194(1):307–322, 2009.

[12] J. Siddique and T. R. Belin. Using an approximate bayesian bootstrap to multiply impute nonignorable missing data. *Computational Statistics & Data Analysis*, 53:405–415, 2008.

[13] E. Triantaphyllou. *Multi-Criteria Decision Making Methods: A Comparative Study*. Kluwer Academic Publishers, Dordrecht/Boston/London, 2000.

ISBN: 978-989-95079-6-8