2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

OWA-Based Fuzzy Rule Interpolation for Group Decision Making

Chengyuan Chen and Qiang Shen
Department of Computer Science
Institute of Mathematics, Physics and Computer Science
Aberystwyth University, UK
Email: {chc16, qqs}@aber.ac.uk

***Abstract*—The goal in group decision making is to ensure that the best decision is made with respect to the available information and knowledge possessed by all group members. However, different types of uncertainty may influence both the assessment of the individual views and the derivation of the overall group-level solution. The difficulty in such decision-making may escalate if the views of all individuals only cover part of the problem space. Systems capable of reasoning through fuzzy interpolation can help. Fuzzy rule interpolation is an important technique for performing inference with sparse rule bases. Even when a given observation has no overlap with the antecedent values of any existing rules, fuzzy rule interpolation may still derive a conclusion. This paper presents an approach for achieving group decision making via fuzzy interpolation. Individual preferences are firstly aggregated by means of a method learned on rough-fuzzy set theory, and rough-fuzzy interpolation is then applied to derive the group-level conclusion. Experimental investigations are carried out and the results are presented to demonstrate the efficacy of the proposed work in guaranteeing the overall decision accuracy.**

I. Introduction

Group decision making (GDM) is a process where a number of individuals attempt to reach a consensus on a certain decision. A group solution is the one that is the most acceptable to all the individuals concerned as a whole. In GDM, both the individuals in the group and the group at large jointly make decisions. To do this, individuals need to express their judgements among a set of alternative opinions. However, different types of uncertainty may influence both the assessment of the individual views and the derivation of the overall group-level solution [1]. These include the following factors: (1) An individual's role (weight) in the generation of the group solutions, since there may be a group leader or leaders who play more important roles in a particular GDM process. (2) An individual's preference for possible decision alternatives, since individuals may have a different understanding for the same information and different experiences in the area of current decision problems. (3) An individual's use of criteria for assessing alternatives, since individuals may often have different judgements in comparing the importance between those criteria. All such types of uncertainty translate into difficulties in determining the final solution by the group. In addition, there are many situations where the potential decision alternatives may be ordered and even depicted on an underlying continuum [2]. Each individual may have an optimum or most preferred position on the continuum. Obviously, the closer any given alternative lies to the optimum, the more it may be preferred over another. Sometimes, an individual's optimum may be located between two distinct alternatives. That is, a different preference may appear beyond given alternatives, leading to the difficulty of making a consensual decision.

It is well-known that human judgement including preferences is often subjective, vague and imprecise. Fuzzy systems play an important role in decision making and offer a flexible framework for GDM. Indeed, fuzzy rules are often employed by human beings to make decisions. Such rules use a series of IF-THEN statements to describe what action should be taken in terms of the currently observed information. They are widely used in fuzzy inference systems (FIS) to perform decision making according to given individuals' preferences.

The compositional rule of inference [3] offers an effective mechanism to deal with fuzzy inference for dense rule bases. Given such a rule base and an observation that is at least partially covered by the rule base, the conclusion can be inferred from certain rules that intersect with the observation. However, for the case where a fuzzy rule base contains 'gaps' (termed: sparse rule base [4]), if a given observation has no overlap with the antecedent values of any rule, conventional fuzzy inference methods cannot derive a conclusion. This is of particular significance when a given preference lies between two known alternatives in GDM. Fortunately, using fuzzy rule interpolation (FRI) [5], [6], certain decisions may still be reached. However, different types of uncertainty may influence both the assessment of the individual views and the derivation of the overall group-level solution in GDM. To cope with such uncertain information and knowledge, higher order fuzzy representation may be helpful [7], [8].

In this paper, an FRI technique for GDM is proposed in order to better address the underlying relative uncertainty, thereby determining appropriate decisions. For each criterion, the OWA operator [9] is employed to decide each individual's role. Then, aggregation of individuals' preferences is performed by means of a rough-fuzzy (RF) set theoretic approach [7]. Finally, a popular transformation-based FRI (abbreviated to T-FRI hereafter) [10], [11] is utilised to enable required interpolative reasoning.

The rest of this paper is structured as follows. Section II reviews the general concepts of RF sets, OWA aggregation and T-FRI. Section III illustrates the proposed approach for GDM, and details its implementation. Section IV provides a simulated example that demonstrates the procedures of the proposed work, and verifies its accuracy in comparison to

978-1-4799-2072-3/14/$31.00

possible alternative techniques. The paper is concluded in Section V, including suggestions for possible further work.

II. Background

A. Rough-Fuzzy Sets

Fuzzy set theory and rough set theory are distinct but complementary extensions of conventional set theory. Typical examples which involve fuzzy and rough sets include those as reported in [12], [13], [14].

Let $I = (\mathbb{U}, \mathbb{A})$ be an information system, where $\mathbb{U}$ is a nonempty set (the universe) of finite objects and $\mathbb{A}$ is a nonempty finite set of attributes such that $a : \mathbb{U} \rightarrow V_a$ for every $a \in \mathbb{A}$ with V_a being the set of values that attribute a may take from. With any $P \subseteq \mathbb{A}$ there is a crisp equivalence relation $IND(P)$ [15]:

$$IND(P) = \{(x, y) \in \mathbb{U}^2 \mid \forall a \in P,\ a(x) = a(y)\} \tag{1}$$

If $(x, y) \in IND(P)$, then x and y are indiscernible by attributes from P. The equivalence classes of the indiscernibility relation with respect to P are denoted $[x]_P$, $x \in \mathbb{U}$.

Let $X \subseteq \mathbb{U}$, X can be approximated using only the information contained within P by constructing the P-*lower* and P-*upper* approximations of X [15]:

$$\begin{aligned} \underline{P}X &= \{x | [x]_P \subseteq X\} \\ \overline{P}X &= \{x | [x]_P \cap X \neq \emptyset\} \end{aligned} \tag{2}$$

The tuple $< \underline{P}X, \overline{P}X >$ is called a rough set.

Definition 2.1. [7] With any $P \subseteq \mathbb{A}$, a new equivalence relation $IND(P)$ can be defined by:

$$IND(P) = \{x \in \mathbb{U}^2 \mid \forall F_l(x) \in P,\ F_l \in C_o\} \tag{3}$$

where F_l, $l \in \{1, \ldots, J\}$, are fuzzy sets that are known to exactly belong to a given concept C_o, with C_o, $o \in \{1, \ldots, Y\}$, being a given decision class in X, i.e., $X = \{C_1, C_2, \ldots, C_Y\}$.

Using this equivalence relation, the lower and upper approximations for each C_o in X can be redefined as follows.

Definition 2.2. [7] Let P be an equivalence relation on X and F_l, $l \in \{1, \ldots, J\}$, be fuzzy sets in C_o ($C_o \in X$), the lower and upper approximations are a pair of fuzzy sets with membership functions defined by the following, respectively:

$$\begin{aligned} \mu_{\underline{P}C_o}(x \in [x]_P) &= \inf\{\mu_{F_l}(x) | x \in [x]_P\},\ l \in \{1, \ldots, J\} \\ \mu_{\overline{P}C_o}(x \in [x]_P) &= \sup\{\mu_{F_l}(x) | x \in [x]_P\},\ l \in \{1, \ldots, J\} \end{aligned} \tag{4}$$

The tuple $< \underline{P}X, \overline{P}X >$ is called an RF set, in contrast to the most general use of this term in the literature [16].

B. OWA Operator

When dealing with real-world problems, the opinions of different experts are usually aggregated in order to provide more robust solutions. This reflects a nature of GDM. Apart from the classical aggregation operators, a family of Ordered Weighted Averaging (OWA) operators [9] have also been successfully applied [17], [18]. The fundamental aspect of an OWA operator is the reordering step in which the inputs are rearranged in descending order and then integrated into a single aggregated value.

Definition 2.3. [9] An OWA operator of dimension p is a mapping $\mathbb{R}^p \rightarrow \mathbb{R}$, which has an associated weighting vector $W = (w_1, w_2, \ldots, w_p)^T$, where $w_v \in [0, 1]$ and $\sum_{v=1}^{p} w_v = 1$. An input vector $(c_1, c_2, \ldots, c_p)$, is aggregated as follows:

$$\text{OWA}(c_1, c_2, \ldots, c_p) = \sum_{v=1}^{p} w_v \hat{c}_v \tag{5}$$

where $\hat{c}_v$ is the vth largest element in the vector $(c_1, c_2, \ldots, c_p)$ and $\hat{c}_1 \geq \hat{c}_2 \geq \cdots \geq \hat{c}_p$.

One important method required to implement any OWA operator is to determine the associated weights. In general, different choices of the weight vector W lead to different aggregation results. Three special instances of OWA are the classical Average, Max and Min. The Average operator results by setting $w_v = 1/p$, the Max by $w_1 = 1$ and $w_v = 0$ for $v \neq 1$, and the Min by $w_p = 1$ and $w_v = 0$ for $v \neq p$. Apart from these, other approaches for obtaining the OWA weights can be classified into two categories, namely: *argument-independent* and *argument-dependent*. As reflected by their respective names, the weights derived by the former are not related to the arguments being aggregated, while the latter determines the weights on the basis of the input arguments.

1) DOWA Operator:

Definition 2.4. [19] Let $(c_1, c_2, \ldots, c_p)$ be the argument vector, and e be the average value of this argument set: $e = \frac{1}{p}\sum_{v=1}^{p} c_v$. The similarity degree between any argument c_v and the average value e is calculated by

$$s(c_v, e) = 1 - \frac{|c_v - e|}{\sum_{u=1}^{p} |c_u - e|} \tag{6}$$

Note that if $\sum_{u=1}^{p} |c_u - e| = 0$, then $c_u - e = 0$, $u \in \{1, \ldots, p\}$. That is, all the values of the arguments are the same. In this case, $s(c_v, e) = 1$, $v \in \{1, \ldots, p\}$.

From this, an input vector $(c_1, c_2, \ldots, c_p)$ can be aggregated by the DOWA operator as follows [19]:

$$\text{DOWA}(c_1, c_2, \ldots, c_p) = \sum_{u=1}^{p} w_u c_u \tag{7}$$

where the weight vector $W = (w_1, w_2, \ldots, w_p)^T$ is generated by

$$w_v = \frac{s(c_v, e)}{\sum_{u=1}^{p} s(c_u, e)},\quad v \in \{1, \ldots, p\} \tag{8}$$

2) Clus-DOWA Operator:

Definition 2.5. [20] Let $(c_1, c_2, \ldots, c_p)$ be an argument vector. For each argument c_u, the concept of its reliability r_u is defined as its distance d_u to the nearest cluster recorded during a given clustering process, i.e.,

$$r_u = 1 - \frac{d_u}{\sum_{v=1}^{p} d_v} \tag{9}$$

Note that if $\sum_{v=1}^{p} d_v = 0$, then $d_v = 0$, $v \in \{1,\ldots,p\}$. This is a similar case to that mentioned previously, therefore $r_u = 1$, $u \in \{1,\ldots,p\}$.

From this, a specific and powerful OWA operator can be defined as follows.

Definition 2.6. [20] The Clus-DOWA operator is defined by

$$\text{Clus-DOWA}(c_1, c_2, \ldots, c_p) = \sum_{v=1}^{p} w_v c_v \tag{10}$$

where the weight vector is calculated from a computed vector of reliability measurement $(r_1, r_2, \ldots, r_p)$:

$$w_v = \frac{r_v}{\sum_{u=1}^{p} r_u}, \quad v \in \{1,\ldots,p\} \tag{11}$$

C. Transformation-Based Fuzzy Rule Interpolation

An outline of the existing T-FRI is provided in this subsection, including both the underlying concepts and the interpolation steps, further details can be found in [10], [11]. For simplicity, only rules involving trapezoidal-shaped membership functions are considered here.

Given a trapezoidal fuzzy set A, denoted by (a_0, a_1, a_2, a_3), where a_0 and a_3 are the two limit points of the support of A where membership values equal 0, a_1 and a_2 are the two limit normal points of A where membership values equal 1. Its *representative value* (Rep) is defined such that:

$$\text{Rep}(A) = \frac{a_0 + \frac{a_1+a_2}{2} + a_3}{3} \tag{12}$$

Note that this definition subsumes the Rep of a triangular fuzzy set as its specific case (where a_1 and a_2 of a trapezoid are collapsed into a single value a_1).

1) Closest N Rules Selection: Without losing generality, suppose that a rule R_i and an observation O are represented by:

R_i : if x_1 is A_{i1}, $\cdots$, x_j is A_{ij}, $\cdots$, x_M is A_{iM}
then y is B_i
O : x_1 is A_1^*, $\cdots$, x_j is A_j^*, $\cdots$, x_M is A_M^*

where A_{ij} denotes the *j*th antecedent fuzzy set of Rule R_i, A_j^* denotes the observed fuzzy set of variable x_j, and B_i denotes the consequent fuzzy set of Rule R_i with $j \in \{1,\ldots,M\}$, M being the number of antecedent variables.

The distances d_{ij} between the pairs of A_{ij} and A_j^* can be calculated as follows:

$$d_{ij} = d(A_{ij}, A_j^*) = d(\text{Rep}(A_{ij}), \text{Rep}(A_j^*)) \tag{13}$$

The distance d_i between the rule R_i and the observation O is deemed to be the average of all antecedent variables' distances:

$$d_i = \sqrt{\sum_{j=1}^{M} {d'_{ij}}^2}, \quad d'_{ij} = \frac{d_{ij}}{\max_j - \min_j} \tag{14}$$

where $\max_j$ and $\min_j$ are the maximum and minimum values of variable j, $j \in \{1,\ldots,M\}$. Each distance measure d_{ij} is normalised into the range $[0,1]$, denoted by d'_{ij}, to make the absolute distances compatible with each other over different domains.

2) Intermediate Rule Construction: Suppose N ($N \geq 2$) closest rules have been chosen from the observation. Such rules are represented as R_i, $i \in \{1,\ldots,N\}$, each has M antecedents A_{ij}, $j \in \{1,\ldots,M\}$. Let $w_{A_{ij}}$ denote the weight to which the *j*th antecedent of the *i*th rule contributes to the intermediate rule. The normalised weight $w'_{A_{ij}}$ can be defined as:

$$w'_{A_{ij}} = \frac{w_{A_{ij}}}{\sum_{i=1}^{N} w_{A_{ij}}}, \quad w_{A_{ij}} = \frac{1}{d_{ij}} \tag{15}$$

Note that if $d_{ij} = 0$, then $\text{Rep}(A_{ij}) = \text{Rep}(A_j^*)$. In this case, the antecedent of the observation is considered to be 'identical' to the corresponding antecedent of the rule R_i, in terms of the currently applied definition of Rep. Thus, $w_{A_{ij}} = 1$ for the 'identical' one(s), while $w_{A_{ij}} = 0$ for the remainder.

The antecedent of the so-called intermediate fuzzy term A_j^{IFT} is constructed from the antecedents of these closest rules. Another process *shift* is then introduced to modify A_j^{IFT} to the antecedent of the intermediate rule A'_j so that it will have the same Rep as A_j^*:

$$A'_j = A_j^{IFT} + \delta_{A_j}(\max_j - \min_j), \quad A_j^{IFT} = \sum_{i=1}^{N} w'_{A_{ij}} A_{ij} \tag{16}$$

where δ_{A_j} is a constant defined by:

$$\delta_{A_j} = \frac{\text{Rep}(A_j^*) - \text{Rep}(A_j^{IFT})}{\max_j - \min_j} \tag{17}$$

Regarding the consequence of the intermediate rule B', it can be calculated by analogy to the computation of the antecedent, such that:

$$B' = B^{IFT} + \delta_B(\max - \min), \quad B^{IFT} = \sum_{i=1}^{N} w'_{B_i} B_i \tag{18}$$

where B^{IFT} is the consequence of the intermediate fuzzy term, max and min are the maximum and minimum values of consequent variable, w'_{B_i} and δ_B are the means of $w'_{A_{ij}}$ and δ_{A_j}, $i \in \{1,\ldots,N\}$, $j \in \{1,\ldots,M\}$, respectively, which are defined as:

$$w'_{B_i} = \frac{1}{M} \sum_{j=1}^{M} w'_{A_{ij}}, \quad \delta_B = \frac{1}{M} \sum_{j=1}^{M} \delta_{A_j} \tag{19}$$

3) Scale Transformation: The similarity degree between two fuzzy sets A' and A^* is measured by two *scale rates* s_b and s_t ($s_b \geq 0$ and $s_t \geq 0$), and one *scale ratio* $\mathbb{S}$.

The scale rates s_b and s_t for scaling the bottom and top supports of A' with respect to A^* are defined by:

$$s_b = \frac{a_3^* - a_0^*}{a'_3 - a'_0}, \quad s_t = \frac{a_2^* - a_1^*}{a'_2 - a'_1} \tag{20}$$

resulting in A''. The scale ratio $\mathbb{S}$, which represents the actual increase of the ratio between the bottom support and the top support, is then introduced to further modify A'' to avoid the top support of the resultant fuzzy set becoming wider than the bottom support, such that:

$$\mathbb{S} = \begin{cases} \dfrac{\frac{a_2^*-a_1^*}{a_3^*-a_0^*} - \frac{a_2'-a_1'}{a_3'-a_0'}}{1 - \frac{a_2'-a_1'}{a_3'-a_0'}} & \text{if } s_t \geq s_b \\ \dfrac{\frac{a_2^*-a_1^*}{a_3^*-a_0^*} - \frac{a_2'-a_1'}{a_3'-a_0'}}{\frac{a_2'-a_1'}{a_3'-a_0'}} & \text{otherwise} \end{cases} \tag{21}$$

Scale transformation is then applied to generate B'' from B' using s_b' and s_t' under the conditions $\mathbb{S}' = \mathbb{S}$ and $s_b' = s_b$. The scale rate s_t' of the top support of B' is calculated such that

$$s_t' = \begin{cases} s_b' * (\mathbb{S} * \dfrac{b_3' - b_0'}{b_2' - b_1'} - \mathbb{S} + 1) & \text{if } s_t \geq s_b \\ s_t & \text{otherwise} \end{cases} \tag{22}$$

4) Move Transformation: The similarity degree is further reinforced by the use of the *move rate* m. By using m, A'' is moved so that the transformed fuzzy set exactly matches the shape of A^*. The move rate m is defined by:

$$m = \begin{cases} \dfrac{a_0^* - a_0''}{\frac{a_1''-a_0''}{3}} & \text{if } a_0^* \geq a_0'' \\ \dfrac{a_0^* - a_0''}{\frac{a_3''-a_2''}{3}} & \text{otherwise} \end{cases} \tag{23}$$

Note that the above scale and move transformations are utilised for one antecedent only. The calculation for M antecedents is a mere repetition, resulting in s_{bj}', s_{tj}' and m_j, $j \in \{1, \ldots, M\}$.

The final interpolated conclusion B^* is estimated by applying the scale and move transformations to B', using s_{B_b}, s_{B_t} and m_B, such that:

$$s_{B_b} = \frac{1}{M}\sum_{j=1}^{M} s_{bj}', \quad s_{B_t} = \frac{1}{M}\sum_{j=1}^{M} s_{tj}', \quad m_B = \frac{1}{M}\sum_{j=1}^{M} m_j \tag{24}$$

where M is the number of antecedent variables.

III. Proposed Rough-Fuzzy Interpolation for Group Decision Making

The objective of *aggregation* is to combine individuals' preferences into an overall aggregated value so that the final decision takes into account all individuals' contributions. Different but similar opinions are usually aggregated to provide more robust solutions. The particular concern of this work is to deal with the situations where conclusions cannot be inferred but may be interpolated when given uncertain observations have no overlap with any rules.

One possible approach is to interpolate all the conclusions separately with respect to each given observation first and then, to derive the final solution by aggregating all the individual conclusions. This approach is hereafter denoted as the IA method, standing for interpolation before aggregation. However, as outlined previously, the first step of interpolation requires the computation of the closest rules from a given rule base. A distance measure needs to be calculated in order to estimate the proximity between each rule antecedent and observation antecedent. This implies a time complexity of $O(xmn)$, where x is the number of observations to be interpolated, m is the number of antecedent variables, and n is the number of fuzzy rules involved in a rule base. An alternative approach creates an artificial observation by aggregating all the observations first and then, to derive the final solution by performing interpolation over this artificial observation. For obvious reasons, this approach is hereafter denoted as the AI method, which has an overall time complexity of $O(mn)$. The reduction in computation complexity is significant, especially when the number of observations becomes large. Consequently, the AI method is employed herein for problem solving while the results are compared to those obtainable by the IA method. The following presents the theoretical framework of the AI approach.

A. Aggregation

In dealing with individuals' preferences, the pessimistic means is to aggregate such preferences by an intersection operation, in order to ensure that all preferences are satisfied. Opposite to this, the optimistic means is to create the artificial overall preference by performing a union operation in an effort to satisfy at least a single preference. To enable the representation of different types of uncertainty, RF sets [7] can be used to support the aggregation. Thus, Definition 2.2 can be applied for situations where all opinions share a common point. Unfortunately, for many instances, individuals may attempt to conceal their preferences for purposes of taking certain strategic advantages or simply misrepresent their own preferences due to lack of sufficient information [2]. This may lead to preferences that are distinct from the others, resulting in an empty intersection (although it will not affect the union). However, all of the individuals should contribute to the outcome, although one outlier should not affect the overall result. This work therefore extends the original definition of RF sets to a more general version with the use of the OWA operators, which is defined as follows.

Definition 3.1. Let P be an equivalence relation on X and F_l, $l \in \{1, \ldots, J\}$, be fuzzy sets in C_o ($C_o \in X$), the lower approximation (LA) and upper approximation (UA) are a pair of fuzzy sets with membership functions defined by the following, respectively:

$$\mu_{\underline{P}C_o}(x \in [x]_P) = \text{OWA}\{\mu_{F_l}(x) | x \in [x]_P\} = \sum_{l=1}^{J} w_l \mu_{F_l}(x)$$
$$\mu_{\overline{P}C_o}(x \in [x]_P) = \text{OWA}\{\mu_{F_l}(x) | x \in [x]_P\} = \sum_{l=1}^{J} w_l \mu_{F_l}(x) \tag{25}$$

where the weight vector $W = (w_1, w_2, \ldots, w_J)^T$ can be computed using different operators as mentioned before.

Note that when using the Min and Max operators for the calculation of LAs and UAs respectively, the results remain the same as those in Definition 2.2. That is, the original is a specific case of this new definition.

The fuzzy sets are aggregated using a partitioning-based method to discretise the input space in this work. The domain of each observed variable x_j, $j = 1, \ldots, M$, is partitioned into a set of discretised values $D_j = \{F_{j1}, \ldots, F_{j|D_j|}\}$, where $|D_j|$ denotes the cardinality of this set. Therefore, given J observations of a variable x_j, the aggregated observation of this variable is calculated using the following OWA operator:

$$F_{\text{OWA}_j} = \sup_{k\in\{0,\ldots,|D_j|\}} \sum_{l=1}^{J} w_{kl}\mu_{F_l}(\min_{x_j} + k * \frac{\max_{x_j} - \min_{x_j}}{|D_j|}) \tag{26}$$

where $\max_{x_j}$ and $\min_{x_j}$ are the maximum and minimum values of the jth observed values F_{jl}, $l = 1, \ldots, J$.

B. Interpolation

The existing T-FRI approach is extended to implement interpolation for RF sets. An RF set A can be represented by the LA A^L and the UA A^U, i.e., $A =< A^L, A^U >$. In particular, when trapezoidal membership functions are used, such an RF set can be illustrated as shown in Fig. 1, where $A^L = (a_0^L, a_1^L, a_2^L, a_3^L; H_{A_1}^L, H_{A_2}^L)$, $A^U = (a_0^U, a_1^U, a_2^U, a_3^U; H_{A_1}^U, H_{A_2}^U)$, with $(a_0^L, a_1^L, a_2^L, a_3^L)$ and $(a_0^U, a_1^U, a_2^U, a_3^U)$ denoting the four limit points of the LA and those of the UA, respectively, and $H_{A_E}^L$ and $H_{A_E}^U$, $E \in \{1, 2\}$, denoting the maximum membership values of A^L and A^U, with $a_0^U \le a_0^L$, $a_3^L \le a_3^U$, $0 < H_{A_1}^L = H_{A_2}^L \le H_{A_1}^U = H_{A_2}^U = 1$. Clearly, the closer the shapes of A^L and A^U are, the less uncertain the information contained within A is. When A^L coincides with A^U, the RF set degenerates to a conventional fuzzy set.

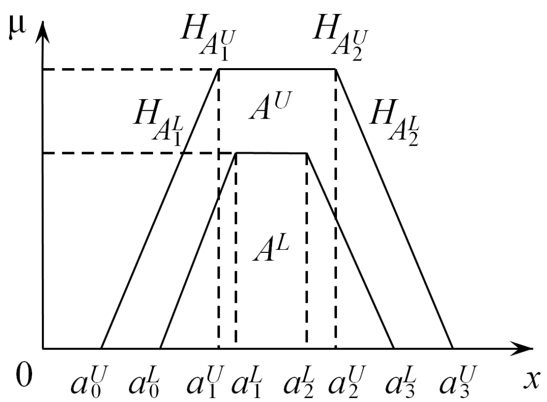

Fig. 1. Lower approximation A^L and upper approximation A^U of a trapezoid RF set A

Suppose that an RF set A as defined in Fig. 1 has the following eight distinct coordinates: $(a_0^L, 0)$, $(a_1^L, H_{A_1}^L)$, $(a_2^L, H_{A_2}^L)$, $(a_3^L, 0)$, $(a_0^U, 0)$, $(a_1^U, H_{A_1}^U)$, $(a_2^U, H_{A_2}^U)$ and $(a_3^U, 0)$. The lower and upper Reps $\text{Rep}(A^L)$ and $\text{Rep}(A^U)$ of A can then be computed according to Eq. (12), such that:

$$\begin{cases} \text{Rep}(A^K)_x = \dfrac{a_0^K + \frac{a_1^K + a_2^K}{2} + a_3^K}{3} \\ \text{Rep}(A^K)_y = \dfrac{0 + \frac{H_{A_1}^K + H_{A_2}^K}{2} + 0}{3} = \dfrac{H_{A_1}^K + H_{A_2}^K}{6} \end{cases} \tag{27}$$

where $K = L, U$, and x and y denote the x coordinate and the y coordinate, respectively.

Note that in the existing T-FRI $\text{Rep}(A)_y$ is a constant, only the x coordinate value is therefore considered there. However, this is no longer the case in this work due to the introduction of higher order uncertainty, both values of x and y coordinates need to be considered. The calculation for $\text{Rep}(A)_y$ follows the same as $\text{Rep}(A)_x$ to maintain consistency.

In order to distinguish different shapes of RF sets, the *shape diversity factor* f needs to be introduced. This work follows the conventional definition of statistical standard deviation (although this may be defined differently).

Definition 3.2. The lower and upper shape diversity factors f_A^L and f_A^U are defined by:

$$f_A^K = \sqrt{\frac{\sum_{g=0}^{2} (\hat{a}_g^K - \text{Rep}(A^K)_x)^2}{3}}, \quad K = L, U \tag{28}$$

where $\hat{a}_0^K = a_0^K$, $\hat{a}_1^K = \frac{1}{2}(a_1^K + a_2^K)$, and $\hat{a}_2^K = a_3^K$.

A small shape diversity factor implies that the four key points of A^L (A^U) tend to be close to those of the lower (upper) Rep. That is, the smaller the shape diversity factor, the smaller the area of the LA (UA).

Extending T-FRI to dealing with RF sets, a single overall Rep of a given RF set is required. For this, the *weight factor* w of the LA (UA) is first introduced below.

Definition 3.3. The lower and upper weight factors w_A^L and w_A^U are defined as the weights of the shape diversity factors, in terms of the areas of the LA and UA, such that:

$$w_A^K = \frac{f_A^K}{f_A^L + f_A^U}, \quad K = L, U \tag{29}$$

where $f_A^L + f_A^U \neq 0$. If however, $f_A^L + f_A^U = 0$, i.e., $f_A^L = 0$ and $f_A^U = 0$, the RF set degenerates to a singleton value, $w_A^L = w_A^U = 1/2$.

Definition 3.4. The overall representative value $\text{Rep}(A)$ of a given RF set A is defined by:

$$\begin{aligned} \text{Rep}(A) =& w_A^L(\text{Rep}(A^L)_x + \text{Rep}(A^L)_y) \\ &+ w_A^U(\text{Rep}(A^U)_x + \text{Rep}(A^U)_y) \end{aligned} \tag{30}$$

where the lower (upper) shape diversity factor is regarded as the weight of the lower (upper) Rep of the LA (UA). This is necessary, as otherwise, the same value for Rep would be derived from different shapes of RF sets.

Given the above definitions, the extended algorithm for deriving the interpolated conclusion with multiple multi-antecedent rules is summarised below.

1) Calculate Representative Values: The lower and upper representative values $\text{Rep}(A^K)_x$ and $\text{Rep}(A^K)_y$ of a given RF set A are calculated first using Eq. (27). The shape diversity factors f_A^K and weight factors w_A^K are computed according to Eqs. (28) and (29), respectively. The overall representative value $\text{Rep}(A)$ is then obtained by Eq. (30), $K = L, U$. The calculations for the antecedents of the observation and all rules follow the same procedure.

2) Choose Closest N Rules: The distances between the artificially created observation and all rules in the rule base are calculated using Eqs. (13) and (14). The N ($N \geq 2$) rules which have minimal distances are then chosen as the N closest rules to perform interpolation.

3) Construct Intermediate Rule: The normalised weight $w'_{A_{ij}}$ of the *j*th antecedent of the *i*th chosen rule, which is calculated by Eq. (15), together with the parameter δ_{A_j}, which is calculated by Eq. (17), are used in Eq. (16) to obtain the antecedent of the intermediate rule A'_j for each antecedent dimension x_j, $i \in \{1, \ldots, N\}$, $j \in \{1, \ldots, M\}$. From this, two parameters w'_{B_i} and δ_B are computed using Eq. (19), and are then used to construct B' from Eq. (18), resulting in the intermediate rule $A'_1 \wedge \cdots A'_j \wedge \cdots A'_M \Rightarrow B'$.

4) Carry out Scale, Move and Height Transformations: In conjunction with the given A^*_j for each antecedent dimension x_j, the rates s^K_{bj}, s^K_{tj}, m^K_j and h_{jE}, $K \in \{L, U\}$, $E \in \{1, 2\}$, can then be calculated using Eqs. (20), (23) and (31). Due to the uncertainty introduced in the membership functions, a further transformation on the heights of the LAs is needed, while the heights of the UAs remain the same owing to its normality. Since the LAs of different RF sets may have different heights, the height transformation is therefore used to transform the heights of A'^L_{jE} to the heights of A^{*L}_{jE}. The *height rate* h is calculated by:

$$h_{jE} = \frac{H^{*L}_{A_{jE}}}{H'^L_{A_{jE}}}, \quad E = 1, 2 \tag{31}$$

where $0 < H^{*L}_{A_{jE}} \leq H^{*U}_{A_{jE}} = 1$ and $0 < H'^L_{A_{jE}} \leq H'^U_{A_{jE}} = 1$, as defined previously. This constraint applies to the interpolated conclusion as well. That is, if the height of B^{*L} is greater than the height of B^{*U} after the height transformation, then $H^{*L}_{B_E} = H^{*U}_{B_E}$.

5) Derive Interpolated Conclusion: The second intermediate term B'' and the interpolated result B^* can then be estimated by the combined $s^K_{B_b}$, $s^K_{B_t}$, m^K_B and h_{B_E}, $K \in \{L, U\}$, $E \in \{1, 2\}$. Here, $s^K_{B_b}$, $s^K_{B_t}$ and m^K_B are computed following Eqs. (21), (22), (23) and (24), respectively, and h_{B_E} is computed according to Eq. (31) such that:

$$h_{B_E} = \frac{1}{M} \sum_{j=1}^{M} h_{jE}, \quad E = 1, 2 \tag{32}$$

Note that no particular information is assumed regarding which preference has a more dominating influence upon the conclusion in general, the arithmetic mean is employed here in order to ensure that the approach is consistent with the underlying T-FRI. This also helps avoid subjective intervention. That is, all the preferences are treated equally. If however, specific information is available, other OWA operators such as the DOWA and Clus-DOWA operators may be more suitable.

6) Implement Modified Procedure: To obtain intuitive interpolated conclusions for RF sets, the relative location between the LA and UA of an RF set should be considered. For this purpose, B'' is modified into B''_c to maintain the relative location both before and after the scale transformation. Here, a *relative location factor* θ is defined by

$$\theta = \frac{B'^L}{B'^U} = \frac{B''^L}{B''^U_n} = \frac{B''^L_n}{B''^U} \tag{33}$$

where B''^L_n and B''^U_n denote the 'new' terms which are modified from the given B''^U and B''^L respectively, using the same θ. The combined B''^L_c and B''^U_c of B''_c are then computed as the mean of the corresponding two terms, such that:

$$B''^K_c = \frac{B''^K + B''^K_n}{2}, \quad K = L, U \tag{34}$$

Similarly, the final interpolated conclusion can also be modified from B^* to B^*_c using the same θ to maintain the relative location both before and after the move transformation.

IV. Experiment and Evaluation

A simulated example is used in this section to validate the efficacy of the proposed work. The results obtainable by the proposed AI method are utilised to compare with those by the two IA methods (the proposed and an existing technique).

A. Experimental Set-up

Individuals may represent their opinions in the form of crisp or fuzzy terms. Occasionally, when only crisp numbers are provided, a fuzzification process is needed. In this simulation-based experimentation, a base function of three crisp input variables, shown in Eq. (35) is chosen to establish a sparse rule base. A fuzzy rule is generated by fuzzifying the crisp inputs and their associated function output, where a numerical value a is converted to a fuzzy set A with a random function f: $A = ((a - f) - f, a - f, a + f, (a + f) + f)$. This provides a simple non-linear (sparse) rule base suitable for the purpose of current investigation.

$$y = 1 + \sqrt{x_1} + \frac{1}{x_2} + \frac{1}{\sqrt{x_3^3}} \tag{35}$$

To evaluate the proposed approach, the output y which is computed from the base function, is assumed to be the ground truth for interpolated results. Without losing generality, the arithmetic mean is used for the OWA operator and regarded as the ground truth for the outcome of the aggregation process.

The first comparison is between the proposed AI and IA methods. In this comparison, the proposed RF sets are applied to aggregate the derived individuals' solutions in IA and the observed opinions in AI, respectively. The Max operator is selected to calculate LAs, while the DOWA and Clus-DOWA operators are used to compute UAs in order to ensure a purely data-driven implementation. For the sake of reducing computational complexity, the aggregated results are simulated with trapezoidal membership functions. The proposed RF interpolation is employed in both IA and AI methods. Thus, two opposite processes are implemented with the proposed approach.

The comparison is also carried out between the proposed AI method and an existing IA method where T-FRI is used for interpolation and the defuzzification-based least squares (D-LS) [21] for aggregation. The weight function used in the existing work of [21] is defined by

$$w_q = \frac{1/\text{Rep}(O_q)}{\sum_{q=1}^{Q} 1/\text{Rep}(O_q)}, \quad q = 1, \ldots, Q \tag{36}$$

where $\text{Rep}(O_q)$ is the Rep of the qth computed output value O_q.

In the present simulation-based experimental evaluation, the Reps of the resultant sets of using IA or AI are recorded. They are then compared against their corresponding ground truth calculated using the base function. The range error (RE) and the root-mean-square error (RMSE) are adopted here to analyse the accuracy of the three different approaches:

$$\begin{aligned} \epsilon_{RE} &= \frac{|O_q - G_q|}{\max_y - \min_y}, \quad q = 1, \ldots, Q \\ \epsilon_{RMSE} &= \sqrt{\frac{1}{Q}\sum_{q=1}^{Q}(O_q - G_q)^2} \end{aligned} \tag{37}$$

where $\max_y$ and $\min_y$ are the maximum and minimum values of the consequent variable, and O_q and G_q denote the qth computed output value and its corresponding ground truth, respectively.

B. Results and Discussion

Since stochastic elements are presented in the generation of observations, the evaluation process is repeated 100 times. Tables I, II and III list the percentage results of the averaged RE and RMSE, where AN is the number of antecedent variables and O is the number of individual observations. The former two tables show the first comparison with the DOWA operator being used in Table I and the Clus-DOWA operator in Table II, while the results of using the proposed AI method that obtained in the first comparison are also utilised for the second comparison as listed in Table III.

It is obvious that for the first comparison, overall, the AI method outperforms the IA method, especially when the number of observations becomes large. The accuracy of the proposed approach is generally higher than that of its opposite process. This is achieved with less computational complexity (as pointed out previously).

Note that the accuracy attainable by the AI method is not so good as its counterpart in the second comparison when the number of observations is small. However, it is important to point out that the computational overheads of IA is significantly greater than that of AI. Thus, IA may be difficult for particular GDM applications with a larger number of opinions or where a timely generation of solutions is required. This is verified by the result in that the accuracy of AI improves and becomes comparable to that of IA as the number of observations is increased. This implies that the proposed approach is suitable for complex systems in GDM. In addition, the accuracy of using the Clus-DOWA operator is consistently (with just one exception) higher than that of utilising the DOWA operator.

V. Conclusion

This paper has presented an OWA-based FRI technique for GDM. The proposed RF set theoretic approach and the extended T-FRI are employed for aggregating individuals' preferences and interpolating the final decision in a purely data-driven manner. According to the simulated experimentation, the proposed technique can reduce the system processing time, while assuring the decision accuracy. This demonstrates that the proposed work is useful for GDM in complex systems.

Although promising, there is still room to improve the current approach. For instance, only centralised clusters are considered for aggregation in this work. It may be beneficial to investigate how the approach would perform when faced with distributed clusters [20]. In this case, a similarity measure would be required. Also, redundant opinions may be removed by classifier ensemble reduction [22] to increase group diversity and produce better results. Additionally, other aggregation operators [23], [24], [25] and FRI methods [26], [27] could also be applied. It would be interesting to compare the results with those. The effect of the proposed technique on the overall efficacy of real-world applications remains active research.

References

[1] J. Lu, G. Zhang, and D. Ruan, "Intelligent multi-criteria fuzzy group decision-making for situation assessments," *Soft Computing*, vol. 12, no. 3, pp. 289–299, 2008.

[2] M. F. Kaplan and C. E. Miller, "Group decision making and normative versus informational influence: effects of type of issue and assigned decision rule," *Journal of Personality and Social Psychology*, vol. 53, no. 2, p. 306, 1987.

[3] L. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, no. 1, pp. 28–44, 1973.

[4] K. Wong, D. Tikk, T. Gedeon, and L. Kóczy, "Fuzzy rule interpolation for multidimensional input spaces with applications: a case study," *IEEE Transactions on Fuzzy Systems*, vol. 13, no. 6, pp. 809–819, 2005.

[5] L. Kóczy and K. Hirota, "Approximate reasoning by linear rule interpolation and general approximation," *International Journal of Approximate Reasoning*, vol. 9, no. 3, pp. 197–225, 1993.

[6] ——, "Interpolative reasoning with insufficient evidence in sparse fuzzy rule bases," *Information Sciences*, vol. 71, no. 1-2, pp. 169–201, 1993.

[7] C. Chen and Q. Shen, "A new method for rule interpolation inspired by rough-fuzzy sets," in *Proceedings of 2012 IEEE International Conference on Fuzzy Systems*, 2012, pp. 1098–1105.

[8] C. Chen, C. Quek, and Q. Shen, "Scale and move transformation-based fuzzy rule interpolation with interval type-2 fuzzy sets," in *Proceedings of 2013 IEEE International Conference on Fuzzy Systems*, 2013, pp. 1–8.

[9] R. R. Yager, "On ordered weighted averaging aggregation operators in multicriteria decisionmaking," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 18, no. 1, pp. 183–190, 1988.

[10] Z. Huang and Q. Shen, "Fuzzy interpolative reasoning via scale and move transformations," *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 2, pp. 340–359, 2006.

[11] ——, "Fuzzy interpolation and extrapolation: a practical approach," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 1, pp. 13–28, 2008.

[12] R. Jensen and Q. Shen, *Computational intelligence and feature selection: rough and fuzzy approaches*. Wiley-IEEE Press, 2008.

[13] ——, "New approaches to fuzzy-rough feature selection," *IEEE Transactions on Fuzzy Systems*, vol. 17, no. 4, pp. 824–838, 2009.

[14] N. Parthaláin, Q. Shen, and R. Jensen, "A distance measure approach to exploring the rough set boundary region for attribute reduction," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 3, pp. 305–317, 2010.

[15] Z. Pawlak, "Rough sets," *International Journal of Computer and Information Sciences*, vol. 11, no. 5, pp. 341–356, 1982.

[16] M. Banerjee and S. Pal, "Roughness of a fuzzy set," *Information Sciences*, vol. 93, no. 3, pp. 235–246, 1996.

[17] G. Bordogna, M. Fedrizzi, and G. Pasi, "A linguistic modeling of consensus in group decision making based on owa operators," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 27, no. 1, pp. 126–133, 1997.

TABLE I. COMPARISON OF ACCURACY - DOWA OPERATOR

	Accuracy	AN = 1			AN = 2			AN = 3		
		O = 5	O = 20	O = 50	O = 5	O = 20	O = 50	O = 5	O = 20	O = 50
IA	$\epsilon_{\%RE}$	1.13	0.87	0.82	1.56	1.59	2.06	1.51	1.18	1.19
	$\epsilon_{\%RMSE}$	10.86	8.41	7.58	15.14	15.22	18.59	14.26	11.34	11.02
AI	$\epsilon_{\%RE}$	1.44	0.80	0.59	1.44	0.80	0.57	1.28	0.98	0.62
	$\epsilon_{\%RMSE}$	12.78	7.49	5.37	14.72	7.83	5.29	12.13	10.15	6.01

TABLE II. COMPARISON OF ACCURACY - CLUS-DOWA OPERATOR

	Accuracy	AN = 1			AN = 2			AN = 3		
		O = 5	O = 20	O = 50	O = 5	O = 20	O = 50	O = 5	O = 20	O = 50
IA	$\epsilon_{\%RE}$	0.85	0.78	0.74	1.15	1.28	1.64	1.00	1.06	1.04
	$\epsilon_{\%RMSE}$	7.85	7.28	6.95	10.77	12.33	15.48	10.15	9.97	9.75
AI	$\epsilon_{\%RE}$	1.01	0.79	0.56	1.27	0.72	0.55	1.19	0.95	0.59
	$\epsilon_{\%RMSE}$	9.72	7.27	5.14	12.85	7.04	5.14	11.04	10.20	5.69

TABLE III. COMPARISON OF ACCURACY - T-FRI AND D-LS

	Accuracy	AN = 1			AN = 2			AN = 3		
		O = 5	O = 20	O = 50	O = 5	O = 20	O = 50	O = 5	O = 20	O = 50
IA	$\epsilon_{\%RE}$	0.71	0.60	0.51	1.10	0.80	0.69	0.79	0.55	0.53
	$\epsilon_{\%RMSE}$	6.45	5.23	4.37	10.32	7.28	5.71	7.44	4.98	4.55
AI (D)	$\epsilon_{\%RE}$	1.44	0.80	0.59	1.44	0.80	0.57	1.28	0.98	0.62
	$\epsilon_{\%RMSE}$	12.78	7.49	5.37	14.72	7.83	5.29	12.13	10.15	6.01
AI (C)	$\epsilon_{\%RE}$	1.01	0.79	0.56	1.27	0.72	0.55	1.19	0.95	0.59
	$\epsilon_{\%RMSE}$	9.72	7.27	5.14	12.85	7.04	5.14	11.04	10.20	5.69

[18] P. Su, C. Shang, and Q. Shen, "Owa aggregation of fuzzy similarity relations for journal ranking," in *Proceedings of 2013 IEEE International Conference on Fuzzy Systems*, 2013, pp. 1–7.

[19] Z. Xu, "Dependent owa operators," in *Modeling Decisions for Artificial Intelligence*. Springer, 2006, pp. 172–178.

[20] T. Boongoen and Q. Shen, "Nearest-neighbor guided evaluation of data reliability and its applications," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 40, no. 6, pp. 1622–1633, 2010.

[21] Y.-M. Wang and C. Parkan, "Two new approaches for assessing the weights of fuzzy opinions in group decision analysis," *Information sciences*, vol. 176, no. 23, pp. 3538–3555, 2006.

[22] R. Diao, F. Chao, T. Peng, N. Snooke, and Q. Shen, "Feature selection inspired classifier ensemble reduction," *IEEE Transactions on Cybernetics*, to appear.

[23] P. Angelov and R. Yager, "Density-based averaging–a new operator for data fusion," *Information Sciences*, vol. 222, pp. 163–174, 2013.

[24] R. R. Yager, "Centered owa operators," *Soft Computing*, vol. 11, no. 7, pp. 631–639, 2007.

[25] R. R. Yager and D. P. Filev, "Induced ordered weighted averaging operators," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 29, no. 2, pp. 141–150, 1999.

[26] L. Yang and Q. Shen, "Adaptive fuzzy interpolation," *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 6, pp. 1107–1126, 2011.

[27] ——, "Closed form fuzzy interpolation," *Fuzzy Sets and Systems*, vol. 225, no. 0, pp. 1–22, 2013.