Backward Fuzzy Rule Interpolation

Shangzhu Jin, Ren Diao, Chai Quek, *Senior Member, IEEE*, and Qiang Shen

***Abstract*—Fuzzy rule interpolation offers a useful means to enhancing the robustness of fuzzy models by making inference possible in sparse rule-based systems. However, in real-world applications of interconnected rule bases, situations may arise when certain crucial antecedents are absent from given observations. If such missing antecedents were involved in the subsequent interpolation process, the final conclusion would not be deducible using conventional means. To address this important issue, a new approach named backward fuzzy rule interpolation and extrapolation (BFRIE) is proposed in this paper, allowing the observations, which directly relate to the conclusion to be inferred or interpolated from the known antecedents and conclusion. This approach supports both backward interpolation and extrapolation which involve multiple fuzzy rules, with each having multiple antecedents. As such, it significantly extends the existing fuzzy rule interpolation techniques. In particular, considering that there may be more than one antecedent value missing in an application problem, two methods are proposed in an attempt to perform backward interpolation with multiple missing antecedent values. Algorithms are given to implement the approaches via the use of the scale and move transformation-based fuzzy interpolation. Experimental studies that are based on a real-world scenario are provided to demonstrate the potential and efficacy of the proposed work.**

***Index Terms*—Backward interpolation, fuzzy rule interpolation (FRI), missing antecedents, transformation-based interpolation.**

I. Introduction

FUZZY rule interpolation (FRI) was originally proposed in [28] and [29]. It is of particular significance for reasoning in the presence of insufficient knowledge or sparse rule bases. When a given observation has no overlap with antecedent values, no rule can be invoked in classical fuzzy inference, and therefore, no consequence can be derived. A number of important interpolation approaches have been proposed in the literature, including [1], [7], [10], [13], [19]–[21], [27], [30], [43], [50]–[53], most of which can be categorized into two classes with several exceptions (e.g., type-II fuzzy interpolation [8], [9], [34]).

The first category of approaches directly interpolates rules whose antecedents match the given observation. The consequence of the interpolated rule is thus the logical outcome. Typical approaches in this group [28], [29], [44] are based on the use of α-cuts ($\alpha \in (0, 1]$). The α-cut of the interpolated consequent fuzzy set is calculated from the α-cuts of the observed antecedent fuzzy sets, and those of all the fuzzy sets that are involved in the rules used for interpolation. Having found the consequent α-cuts for all $\alpha \in (0, 1]$, the consequent fuzzy set is then assembled by applying the Resolution principle.

Manuscript received August 9, 2013; revised November 3, 2013; accepted December 25, 2013. Date of publication January 29, 2014; date of current version November 25, 2014.

S. Jin, R. Diao, and Q. Shen are with the Department of Computer Science, Institute of Mathematics, Physics and Computer Science, Aberystwyth University, Aberystwyth, SY23 3DB, U.K. (e-mail: shj1@aber.ac.uk; rrd09@aber.ac.uk; qqs@aber.ac.uk).

C. Quek is with the School of Computer Engineering, Nanyang Technological University, Singapore 637457 (e-mail: ashcquek@ntu.edu.sg).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TFUZZ.2014.2303474

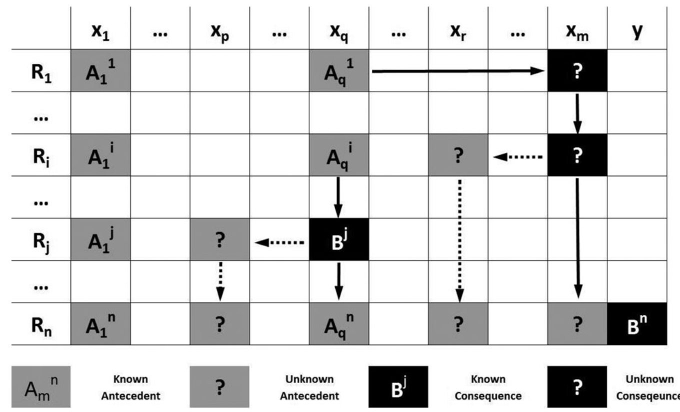


Fig. 1. System structure that may benefit from BFRIE.

The second category is based on the analogical reasoning mechanism [6]. Such approaches first interpolate an artificially created intermediate rule so that the antecedents of the intermediate rule are similar to the given observation [1]. Then, a conclusion can be deduced by firing this intermediate rule through analogical reasoning. The shape distinguishability between the resulting fuzzy set and the consequence of the intermediate rule is then analogous to the shape distinguishability between the observation and the antecedent of the created intermediate rule. In particular, the scale and move transformation-based approach (T-FIR) [20], [21] offers a flexible means to handle both interpolation and extrapolation involving multiple multiantecedent rules.

Despite the numerous approaches developed, FRI techniques are relatively rarely applied in practice [32]. One of the main reasons for this is that many applications involve multiple-input and multiple-output problems. The rules are typically irregular in nature (i.e., not always addressing the same antecedents). In particular, rules may be arranged in an interconnected mesh, where observations and conclusions in between different subsets of rules could be overlapped, and yet not directly related throughout the entire rule base. For such complex systems, any missing values in a given set of observations may lead to failure in interpolation. In Fig. 1, $R_i, i = 1, \ldots, n$ form the rule base, including interpolated rules, and $x_p, x_q, p, q = 1, \ldots, m$ are the variables covering antecedents and consequence. $A_q^i (q = 1, \ldots, m, i = 1, \ldots, n)$ is the fuzzy set on the qth dimension, which is included in the ith rule. The final conclusion B^n of rule R_n cannot be interpolated straightforwardly, because the

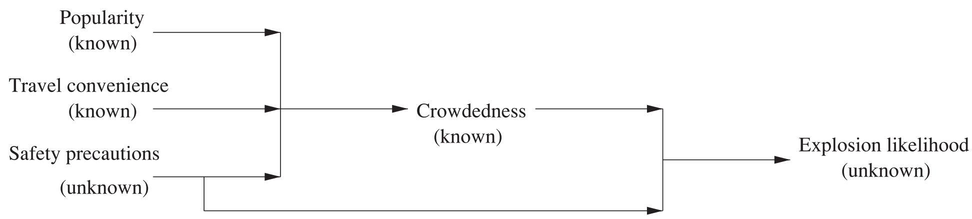


Fig. 2. Hierarchical fuzzy reasoning structure for terrorist bombing threat.

three missing observations A_p^n, A_r^n, and A_m^n cannot be deduced by conventional means.

For instance, consider a practical scenario in detecting terrorist bombing threats. The *Explosion likelihood* may be directly related to the *Crowdedness* of a place and the *Safety precautions*. The number of people in an area may be affected by the *Popularity* of the place, the level of *Travel convenience*, and the amount of *Safety precautions*. A hierarchical structure for this scenario is shown in Fig. 2. For traditional forward interpolative reasoning, in order to interpolate *Explosion likelihood*, the observed values for *Crowdedness* and *Safety precautions* must be both provided. The variable *Safety precautions* is particularly important, as without it, no matter what other information is available, forward interpolation would still fail. Therefore, the interpolation of such crucial missing values may become necessary, in order to allow required inferences to be performed.

To address such problems, this paper proposes a novel approach termed backward fuzzy rule interpolation (BFRIE) by substantially expanding and refining the initial preliminary work of [23] and [24]. This approach enables unknown antecedent values to be interpolated, given other antecedents and the conclusion. Using the earlier example of Fig. 1, the unknown antecedents A_p^n and A_r^n can be backward interpolated according to rules R_j and R_i, where the conclusions B^j, B^i, and the other terms are known. The last missing antecedent value A_m^n can then be interpolated using R_1, and subsequently, B^n can also be computed, as now all required antecedents are known to perform forward interpolation. As such, the proposed techniques support flexible interpolation when certain antecedents are missing from the observation, where traditional FRI methods fail. In addition, BFRIE also enables indirect interpolative reasoning, which involves several fuzzy rules, each with multiple antecedents. Therefore, it offers a means to broaden the application of FRI and fuzzy inference.

General BFRIE (with multiple missing antecedent values) is common in practical problems such as medical diagnosis [16], network intrusion detection [42], oil exploration [49], and intelligence data analysis [4]. To address this challenging issue, two methods are developed. The first directly extends the single missing antecedent case, by computing and searching for good quality parameter combinations for the T-FIR process. The second approach works more closely with conventional FRI procedures by estimating the possible missing antecedent values and, subsequently, verifying the interpolative outcome against the observation.

The remainder of this paper is organized as follows. Section II reviews the general concepts of T-FIR, which is adopted to implement the subsequent developments. Section III introduces the basic form of BFRIE that deals with one single missing antecedent value, along with worked examples. Section IV presents two possible extensions that support the scenarios where multiple antecedent values are missing. Section V describes a real-world application to demonstrate the efficacy of the proposed approach. Systematic randomized experiments are also conducted in order to better compare and verify the accuracies of the proposed methods. Section VI concludes the paper and suggests future enhancements.

II. Background of Transformation-Based Interpolative Reasoning

This section introduces the interpolation procedures involved in T-FIR [21], and defines its underlying key concepts. T-FIR offers a flexible means to handling both interpolation and extrapolation involving multiple multiantecedent fuzzy rules. It guarantees the uniqueness, normality, and convexity of the resulting fuzzy sets. It is also able to handle various fuzzy set representations, including polygonal, Gaussian, and bell-shaped fuzzy membership functions. However, triangular and trapezoidal membership functions are the most frequently used fuzzy set representations in fuzzy systems. Therefore, they are adopted in the algorithm description below.

A key concept used in T-FIR is the representative value $\mathrm{Rep}(A)$ of a given fuzzy set A. When trapezoidal representation is used, $\mathrm{Rep}(A)$ is defined as the center of gravity of its four points (a_0, a_1, a_2, a_3)

$$\mathrm{Rep}(A) = \frac{a_0 + \frac{a_1+a_2}{2} + a_3}{3} \tag{1}$$

where a_0, a_3 represent the left and right extremities (with membership values 0), and a_1, a_2 denote the normal points (with membership value 1), as shown in Fig. 3(a).

As a specific case of trapezoids, where a_1 and a_2 are collapsed into a single value a_1, the fuzzy set becomes a triangular set (a_0, a_1, a_2). In particular, the corresponding $\mathrm{Rep}(A)$ degenerates to the average value of the triple, as given below and shown in Fig. 3(b)

$$\mathrm{Rep}(A) = \frac{a_0 + a_1 + a_2}{3}. \tag{2}$$

In the following, the T-FIR method is outlined using trapezoidal fuzzy sets unless otherwise stated.

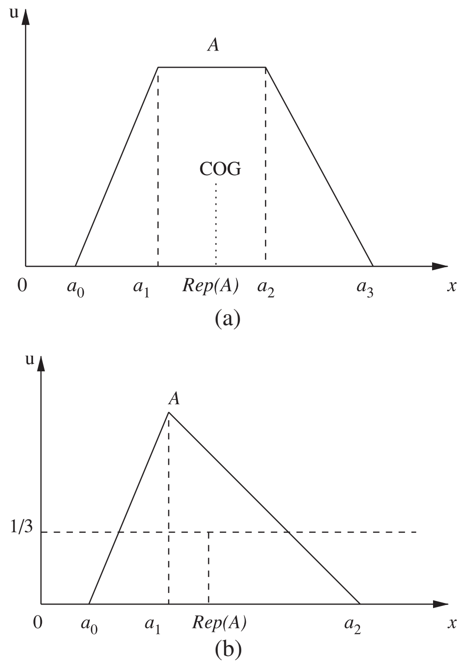


Fig. 3. Representative values of fuzzy sets. (a) Trapezoidal. (b) Triangular.

A. Determination of the Closest Rules

In this paper, given a rule base $\mathbb{U}$, a fuzzy rule $R \in \mathbb{U}$ with M antecedents $A_k, k = 1, 2, \ldots, M$, and an observation O are expressed in the following format:

R: IF x_1 is A_1, …, and x_k is A_k, …, and x_M is A_M, THEN y is B

O: A_1^*, …, A_k^*, …, A_M^*.

The distance d between a rule and an observation is determined by computing the aggregated distance of all the antecedent variables

$$d = \sqrt{\sum_{k=1}^{M} d(A_k, A_k^*)^2}, d(A_k, A_k^*) = \frac{d(\text{Rep}(A_k), \text{Rep}(A_k^*))}{\text{range}_{A_k}} \tag{3}$$

where $\text{range}_{A_k} = \max_{A_k} - \min_{A_k}$ is the domain range of the variable x_k. $d(A_k, A_k^*) \in [0, 1]$ is the normalized result of the otherwise absolute distance measure so that distances are compatible with each other over different variable domains. The N ($N \geq 2$) rules which have the least distance measurements with regard to the observed values A_k^*, and the conclusion B^*, are then chosen to be used in the later steps. To help explain, assume that the observation O and a certain set of closest rules $R_i, i = 1, \ldots, N, R_i \in \mathbb{U}$ that are returned by this step are represented as follows:

O: A_1^*, …, A_k^*, …, A_M^*

R_i: IF x_1 is A_1^i, …, and x_k is A_k^i, …, and x_M is A_M^i, THEN y is B^i.

B. Construction of the Intermediate Rule

Let the normalized displacement factor $\omega_{A_k^i}$ denote the weight of the kth antecedent of R_i

$$\omega_{A_k^i} = \frac{\omega'_{A_k^i}}{\sum_{i=1}^{N} \omega'_{A_k^i}}, \qquad \omega'_{A_k^i} = 1/d(A_k^i, A_k^*). \tag{4}$$

The intermediate fuzzy terms $A_k^\dagger$ that are to be used to build the required intermediate rule are constructed from the antecedents of the N closest rules. These are then shifted to A'_k such that they have the same representative values as those of A_k^*

$$A'_k = A_k^\dagger + \delta_{A_k} \text{range}_{A_k}, \qquad A_k^\dagger = \sum_{i=1}^{N} \omega_{A_k^i} A_k^i \tag{5}$$

where the coordinates of the new fuzzy set $A_k^\dagger$ are calculated on a point-by-point basis, and δ_{A_k} is the bias between A_k^* and $A_k{}'$ on the kth variable domain

$$\delta_{A_k} = d(A_k^*, A_k^\dagger). \tag{6}$$

Similar to (5), the shifted intermediate consequence B' can be computed, with the parameters ω_{B^i} and δ_B being aggregated from the corresponding values of $A_k{}'$, such that

$$B' = \sum_{i=1}^{N} \omega_{B^i} B^i + \delta_B \text{ range}_B$$

$$\omega_{B^i} = \frac{1}{M} \sum_{k=1}^{M} \omega_{A_k^i}, \quad \delta_B = \frac{1}{M} \sum_{k=1}^{M} \delta_{A_k}. \tag{7}$$

C. Scale Transformation

For each antecedent variable of the N chosen rules, the scale transformation works by calculating two scale rates $\overline{s}_{A_k}$ and $\underline{s}_{A_k}$. The support (a'_0, a'_3) of the corresponding shifted fuzzy set A' is transformed into a new support (a''_0, a''_3), and the core (a'_1, a'_2) is transformed into another (a''_1, a''_2), such that

$$\underline{s}_{A_k} = \frac{a''_3 - a''_0}{a'_3 - a'_0} \tag{8}$$

and

$$\overline{s}_{A_k} = \frac{a''_2 - a''_1}{a'_2 - a'_1}. \tag{9}$$

This leads to a scaled fuzzy set $A''_k = (a''_0, a''_1, a''_2, a''_3)$. The corresponding parameters $\underline{s}_B$ and $\overline{s}_B$ of fuzzy set B^* can be calculated as follows:

$$\underline{s}_B = \frac{1}{M} \sum_{k=1}^{M} \underline{s}_{A_k} \qquad \overline{s}_B = \frac{1}{M} \sum_{k=1}^{M} \overline{s}_{A_k}. \tag{10}$$

To maintain the convexity of a scaled fuzzy set, it is necessary to ensure that the scaled support is wider than the core. For this, the following scale ratio $\mathbb{S}$ is applied, which represents the actual increase of the ratios between the core and the support

$$\mathbb{S} = \begin{cases} \dfrac{\frac{a'_2 - a'_1}{a'_3 - a'_0} - \frac{a_2 - a_1}{a_3 - a_0}}{1 - \frac{a'_2 - a'_1}{a'_3 - a'_0}}, & \text{if } \overline{s} \geq \underline{s} \geq 0,\ \mathbb{S} \in [0, 1] \\[2ex] \dfrac{\frac{a'_2 - a'_1}{a'_3 - a'_0} - \frac{a_2 - a_1}{a_3 - a_0}}{\frac{a'_2 - a'_1}{a'_3 - a'_0}}, & \text{if } \underline{s} \geq \overline{s} \geq 0,\ \mathbb{S} \in [-1, 0]. \end{cases} \tag{11}$$

Then, the $\overline{s}_B$ of consequence B^* is relevant to scale ratio $\mathbb{S}$

$$\overline{s}_B = \begin{cases} \frac{\underline{s}_B \mathbb{S}}{\overline{s}_B} - \underline{s}_B \mathbb{S} + \underline{s}_B, & \text{if } \overline{s}_B \geq \underline{s}_B \geq 0 \\ \underline{s}_B \mathbb{S}, & \text{if } \underline{s}_B \geq \overline{s}_B \geq 0. \end{cases} \quad (12)$$

Note that for triangular fuzzy sets, the support (a'_0, a'_2) of the shifted fuzzy set A' is transformed into a new support (a''_0, a''_2), such that the scale rate s_{A_k} is calculated as follows:

$$s_{A_k} = \frac{a''_2 - a''_0}{a'_2 - a'_0}. \quad (13)$$

D. Move Transformation

In general, for multiple antecedent rules, each variable dimension has its own move rate m_{A_k}, in order to move each of the scaled fuzzy sets A''_k to new locations that coincide with those of the originally observed values. This allows the initially constructed intermediate fuzzy terms to completely transform. The final transformed fuzzy sets then match the exact shapes of the observed values A^*_k. Without losing generality, for a given scaled intermediate fuzzy term $A''_k = (a''_0, a''_1, a''_2, a''_3)$, its current support (a''_0, a''_3), and core (a''_1, a''_2) can be moved to (a_0, a_3) and (a_1, a_2), using a move rate m_{A_k} that is calculated as follows:

$$\begin{cases} m_{A_k} = \frac{3(a_0 - a''_0)}{a''_1 - a''_0}, & a_0 \geq a''_0 \\ m_{A_k} = \frac{3(a_0 - a''_0)}{a''_3 - a''_2}, & \text{otherwise.} \end{cases} \quad (14)$$

Similar to the scale transformation, the move rate m_B for the consequent dimension can be calculated by obtaining the arithmetic average of those of the antecedent variables, such that

$$m_B = \frac{1}{M} \sum_{k=1}^{M} m_{A_k}. \quad (15)$$

The final interpolated result B^* can now be computed by applying the scale and move transformation to B', using the resulting parameters $\underline{s}_B$, $\overline{s}_B$, and m_B. Note that for triangular fuzzy sets, obviously, the right and center points a''_2 and a''_1 are used (instead of a''_3 and a''_2), when computing the move ratio according to (14), in the case of $a_0 \leq a''_0$.

III. Backward Fuzzy Rule Interpolation and Extrapolation With Single Missing Antecedent Value

BFRIE with single missing antecedent value (S-BFRIE) is proposed for interpolation involving situations where the consequent value is known and the values of all but one antecedent variable are also given. The task is to estimate the value of that single unknown antecedent. Without losing generality, suppose that a conventional FRI is represented as follows:

$$B^* = f_{\text{FRIE}}((A^*_1, \ldots, A^*_l, \ldots, A^*_M), (R_i, i = 1, \ldots, N)) \quad (16)$$

where f_{FRIE} denotes the interpolation/extrapolation process from M observed values, using a set of selected rules $R_i, i = 1, \ldots, N$, that are closest to $\{A^*_l | l = 1, 2, \ldots, M\}$, and B^* is the interpolated conclusion. S-BFRIE can then be defined in the following form:

$$A^*_l = f_{\text{S-BFRIE}}((B^*, A^*_1, \ldots, A^*_{l-1}, A^*_{l+1}, \ldots, A^*_M) \\ (R_i, i = 1, \ldots, N)) \quad (17)$$

where $f_{\text{S-BFRIE}}$ denotes the entire process of obtaining A^*_l, the unknown (or required) observation, which is to be backward interpolated. It uses the N closest rules, with regard to the observed (or predicted) values from the $(M - 1)$ antecedents and the conclusion B^*.

A. Proposed Approach

A close examination of the T-FIR algorithm reveals that, in order to successfully backward interpolate the missing value, a number of closest rules need to be identified first. All of the parameters that are involved in T-FIR (for trapezoidal fuzzy sets) ω, δ, $\underline{s}$, $\overline{s}$, $\mathbb{S}$, and m also need to be computed for the known antecedent variables, and now observed consequent variable. The acquisition of these essential parameters allows a possible transformation process to be derived, which then helps to restore the missing antecedent value. The proposed S-BFRIE algorithm that reflects this intuition is summarized below.

1) Determination of the Closest Rules: In reference to the earlier definition of the S-BFRIE process in (17), when $B^*, (A^*_1, \ldots, A^*_{l-1}, A^*_{l+1}, \ldots, A^*_M)$ are given, in order to interpolate/extrapolate the unknown antecedent A^*_l, the discovery of the closest rules $R_i, i = 1, \ldots, N$, are required. Instead of using the distance measure that is introduced in (3), a modified scheme is proposed in order to reflect the biased consideration toward the consequent variable (as per the intuition indicated previously)

$$d = \sqrt{w_B {d_B}^2 + \sum_{k=1, k \neq l}^{M} (w_{A_k} {d_{A_k}}^2)}. \quad (18)$$

In implementing S-BFRIE, without sufficient expert knowledge on the relative level of significance of different antecedents, all antecedents are treated equally

$$w_B = \sum_{k=1}^{M} w_{A_k} = 1, \; w_{A_1} = w_{A_k} = w_{A_M} = \frac{1}{M}. \quad (19)$$

Note that in choosing the closest rules, the square root used in the original distance measure becomes unnecessary, as only the ordering information is needed. Therefore, the distance calculation can be simplified to

$$\hat{d} = {d_B}^2 + \frac{1}{M} \sum_{k=1, k \neq l}^{M} {d_{A_k}}^2. \quad (20)$$

2) Construction of the Intermediate Fuzzy Terms: To help explain, assume a certain set of closest rules $R_i, i = 1, \ldots, N, R_i \in \mathbb{U}$ that are returned by the previous distance calculation. Following the original T-FIR algorithm, in order to create the intermediate (shifted) fuzzy terms for the known antecedent variables: A'_k, $k = 1, \ldots, M, k \neq l$, the following parameters $w_{A^i_k}$, $i = 1, \ldots, N$, and δ_{A_k} need to be computed first according to (4)–(6). The parameter values for the intermediate (shifted) consequent fuzzy term B': w_{B^i}, $i = 1, \ldots, N$, and δ_B can be computed using exactly the same formulae as those of A_k, since its value B^* is also directly observed.

The formulae given in (7), although no longer needed in this scenario, reveal that both w_{B^i} and δ_B are algebraic averages of the parameter values from individual antecedent terms. For instance, if A_l were not missing, $\omega_{A_l^i}$ would become part of the sum: $\omega_{B^i} = \frac{1}{M}\sum_{k=1}^{M}\omega_{A_k^i}$ in (4). Thus, it has an intuitive appeal to assume that, when backward interpolating a certain parameter value for A_l, say $\omega_{A_l^i}$, the parameter value that is associated with the consequent variable ω_{B^i} should be treated with a biased weight, which is the sum of all antecedent weights. The parameter values for the missing antecedent such as $\omega_{A_l^i}$ are then calculated by subtracting those of the known antecedents from that of the consequent as follows:

$$\omega_{A_l^i} = M\omega_{B^i} - \sum_{k=1,k\neq l}^{M}\omega_{A_k^i}. \tag{21}$$

Following the same logic, ω_{A_l} can be obtained

$$\delta_{A_l} = M\delta_B - \sum_{k=1,k\neq l}^{M}\delta_{A_k}. \tag{22}$$

The acquisition of these parameter values allow the construction of the intermediate (shifted) fuzzy term A'_l for the missing antecedent dimension, similar to (5) and (7)

$$A'_l = A_l^\dagger + \delta_{A_l}\text{range}_{A_l}, \quad A_l^\dagger = \sum_{i=1}^{N}\omega_{A_l^i}A_l^i. \tag{23}$$

Note that according to the characteristics of the T-FIR algorithm, this shifted fuzzy term A'_l also determines the representative value of the final interpolation output A_l^*, since the later transformations will not alter $\text{Rep}(A'_l)$.

3) Scale and Move Transformation: Having obtained the intermediate (shifted) fuzzy terms, the essential parameters $\underline{s}_{A_l}$, $\overline{s}_{A_l}$ (or a single scale rate s_{A_k} for triangular representation), and m_{A_l} that are involved in the transformation process can be derived. Following the same intuition and computational steps as those for $w_{A_k^i}$, $i = 1, \ldots, N$, and δ_{A_l}, by reversing the forward transformation procedure that is introduced in (10) and (15), the required values can be found as follows:

$$\underline{s}_{A_l} = M\underline{s}_B - \sum_{k=1,k\neq l}^{M}\underline{s}_{A_k} \tag{24}$$

$$\overline{s}_{A_l} = M\overline{s}_B - \sum_{k=1,k\neq l}^{M}\overline{s}_{A_k} \tag{25}$$

$$m_{A_l} = Mm_B - \sum_{k=1,k\neq l}^{M}m_{A_k} \tag{26}$$

where $\underline{s}_B$, $\overline{s}_B$, and m_B are immediately obtainable by resolving (8), (9), and (14). Note that to guarantee the transformed fuzzy sets to be convex, $\overline{s}_{A_l}$ should be fixed in terms of the scale ratio $\mathbb{S}_{A_l}$:

$$\mathbb{S}_{A_l} = M\mathbb{S}_B - \sum_{k=1,k\neq l}^{M}\mathbb{S}_{A_k} \tag{27}$$

where $\mathbb{S}_{A_l}$ is the fixed scale ratio of A_l, $\mathbb{S}_B$ is the scale ratio of consequent dimension B, and $\mathbb{S}_{A_k}$ is the scale ratio of A_k, $k = 1, 2, \ldots, M, k \neq l$

$$\overline{s}_{A_l} = \begin{cases} \dfrac{\underline{s}_{A_l}\mathbb{S}_{A_l}}{\overline{s}_{A_l}} - \underline{s}_{A_l}\mathbb{S}_{A_l} + \underline{s}_{A_l}, & \text{if } \overline{s}_{A_l} \geq \underline{s}_{A_l} \geq 0 \\ \overline{s}_{A_l}\mathbb{S}_{A_l}, & \text{if } \underline{s}_{A_l} \geq \overline{s}_{A_l} \geq 0. \end{cases} \tag{28}$$

Finally with all parameters acquired, the transformation on A'_l can be performed, resulting in the (backward) interpolated value A_l^*

$$T(A'_l, A_l^*) = \{\underline{s}_{A_l}, \overline{s}_{A_l}, \mathbb{S}_{A_l}, m_{A_l}\}. \tag{29}$$

B. Worked Examples

This section provides three worked examples of the proposed BFRIE approach. For each of these, the value of the consequent variable is obtained by utilizing the T-FIR method (following the forward FRI procedure of [21]), using randomly chosen values for the antecedent variables. The "missing" value is then (purposefully) removed from the observation, allowing the application of BFRIE. The aim of running these examples is twofolded: 1) to demonstrate the correctness of the BFRIE method, i.e., the proposed procedure can indeed restore the originally observed value (with an acceptable degree of error), and 2) to show that the proposed distance measure is effective in identifying relevant rules in order to perform interpolation (noting that the rules that are involved in the initial generation process may, or may not be selected).

Example 3.1: S-BFRIE With Trapezoidal Fuzzy Sets

This example illustrates S-BFRIE involving multiple multi-antecedent rules, where the variable values are represented by trapezoidal membership functions. The observation and the four closest rules are given in Table I and Fig. 4 (while the subprocess of selecting the closet rules is omitted because it is a straightforward application of (18) to the sparse rule base). Here, A_3^* is the missing antecedent, which is to be inferred.

1) *Construction of the Intermediate Fuzzy Terms:* As explained in Section III-A2, the normalized weights of the antecedents and observed conclusion are derived according to (4), their values are shown in Table II. The parameters for the missing observation $\omega_{A_3^i}$, $i = 1, 2, 3, 4$, can then be calculated using (21), resulting in: $\omega_{A_3^1} = 0.04$, $\omega_{A_3^2} = 0.70$, $\omega_{A_3^3} = 0.06$, $\omega_{A_3^4} = 0.19$. From this, the intermediate fuzzy set $A_3^\dagger = (3.39, 4.43, 5.14, 5.60)$ can be obtained according to (22). Then, the bias δ_{A_3} between A_3^* and $A_3^\dagger$ is calculated using (22), which has a value very close to 0 for this particular case, indicating that no further shifting is necessary. Therefore, the value of the shifted fuzzy term $A'_3 = (4.19, 5.21, 5.90, 6.49)$ can be obtained from (5), which has the same representative value as A_3^*.

TABLE I
FOUR CLOSEST RULES FOR OBSERVATION

	O	R_1	R_2	R_3	R_4
x_1	$(3.5, 4.0, 5.0, 7.0)$	$(0.2, 1.1, 2.2, 2.7)$	$(2.0, 2.3, 2.5, 3.4)$	$(8.2, 9.5, 10.5, 11.0)$	$(10.5, 11.5, 12.5, 13.1)$
x_2	$(5.0, 5.5, 6.0, 7.5)$	$(1.5, 2.0, 2.5, 3.0)$	$(3.1, 3.2, 3.5, 4.3)$	$(7.5, 9.0, 10.2, 11.3)$	$(10.0, 11.2, 12.3, 13.0)$
x_3	missing	$(0.4, 1.5, 2.0, 2.5)$	$(2.5, 3.5, 4.2, 4.5)$	$(7.3, 9.2, 10.5, 11.1)$	$(10.2, 11.0, 11.5, 13.2)$
x_4	$(4.5, 5.2, 6.5, 7.5)$	$(1.1, 1.5, 2.1, 2.5)$	$(6.1, 7.0, 8.0, 8.6)$	$(3.8, 4.1, 4.3, 5.0)$	$(10.1, 12.0, 12.5, 14.3)$
y	$(5.5, 6.5, 7.0, 8.7)$	$(0.2, 2.0, 2.5, 3.0)$	$(4.0, 4.8, 5.3, 6.0)$	$(9.5, 10.0, 11.3, 12.5)$	$(12.0, 13.0, 13.5, 14.2)$

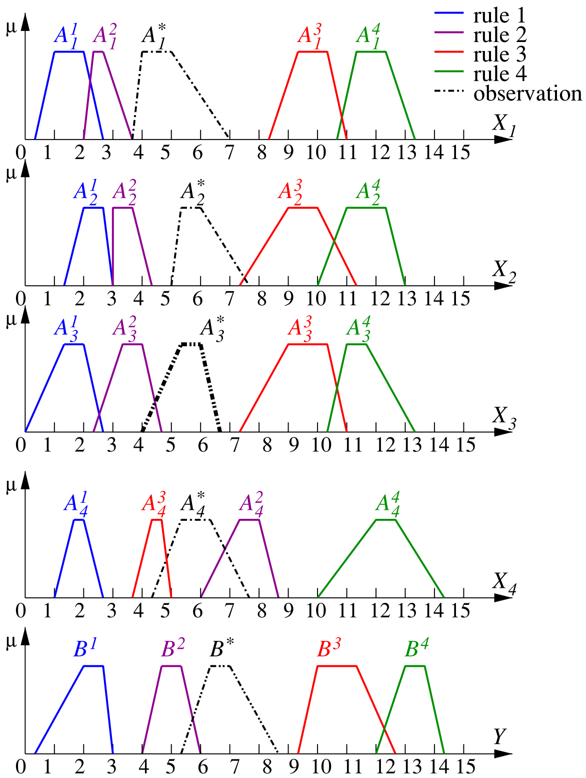


Fig. 4. Example of *B-FRI* with multiple antecedents.

TABLE II
NORMALIZED WEIGHTS OF GIVEN ANTECEDENTS

	R_1	R_2	R_3	R_4
B	0.17	0.45	0.23	0.15
A_1	0.27	0.39	0.20	0.14
A_2	0.23	0.35	0.26	0.16
A_4	0.15	0.36	0.40	0.09

2) *Scale and Move Transformation from A_3' to A_3^*:* The individual scale and move parameters are calculated according to (8)–(15), resulting in $\underline{s}_B = 1.34$, $\overline{s}_B = 0.71$, $m_B = 0.32$. The scale ratio $\mathbb{S}_{B^*} = 0.47$ is obtained using a formula similar to (27). Similarly, the relevant parameters $\underline{s}_{A_k}, \overline{s}_{A_k}, m_{A_k}$ of antecedents A_1^*, A_2^*, A_4^* can be obtained. Following this and using (24)–(27), it can be calculated that $\underline{s}_{A_3} = 1.08$, $\overline{s}_{A_3} = 0.76$, $m_{A_3} = -0.28$, and $\mathbb{S}_{A_3} = 0.70$. The scaled fuzzy term A_3'' is then computed to be $(4.07, 5.32, 5.84, 6.57)$. Finally, the transformed $A_3^* = (4.01, 5.46, 5.98, 6.50)$ can be obtained, which is the estimated missing value for x_3.

TABLE III
TWO CLOSEST RULES FOR OBSERVATION

	O	R_1	R_2
x_1	$(4, 5, 6)$	$(2, 2, 2)$	$(7, 9, 10)$
x_2	$(5, 6, 7)$	$(3, 3, 3)$	$(8, 9, 10)$
x_3	missing	$(4, 4, 4)$	$(9, 10, 11)$
y	$(10, 11, 13)$	$(7, 7, 7)$	$(15, 17, 19)$

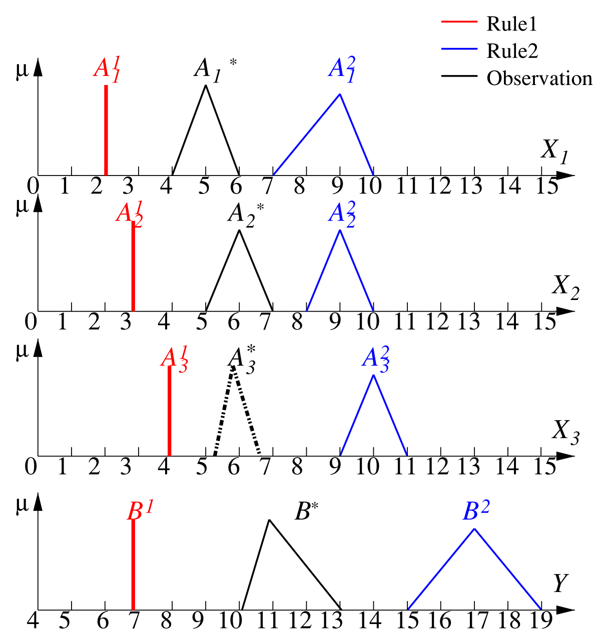


Fig. 5. Example of S-BFRIE with triangular and singleton fuzzy sets.

3) *Verification:* The result of BFRIE can be verified by performing the conventional T-FIR, using the reconstructed observation involving A_3^*. Applying forward interpolation results in the conclusion, $B^* = (5.46, 6.51, 6.85, 8.71)$, $\text{Rep}(B^*) = 6.95$. This is consistent with the given observed conclusion $(5.50, 6.50, 7.00, 8.70)$, which has a representative value of 6.98.

Example 3.2: S-BFRIE With Triangular Fuzzy Sets and Singleton values

To further demonstrate the generality of the proposed approach, this example illustrates S-BFRIE that involves multiple antecedent variables with triangular membership functions and singleton values. The two adjacent rules, which involve singleton fuzzy sets, are given in Table III and Fig. 5, with the observation being $A_1^* = (4, 5, 6)$, $A_2^* = (5, 6, 7)$, $B^* = (10, 11, 13)$.

TABLE IV
THREE CLOSEST RULES FOR OBSERVATION

	O	R_1	R_2	R_3
x_1	(0, 1, 2, 3)	(3.5, 5, 6, 7)	(11, 12, 13, 14)	(7.5, 8, 9, 10)
x_2	missing	(7, 8, 9, 10)	(4, 5, 6, 7)	(10, 11, 12, 13)
x_3	(0, 0.5, 1.5, 2.5)	(10, 11, 12, 13)	(7, 8, 9, 11)	(3, 4, 5, 6)
y	(2.0, 2.9, 3.8, 4.5)	(4.6, 5.8, 6.8, 7.9)	(8, 9, 10, 11)	(11, 12, 13, 14)

1) *Construction of the Intermediate Fuzzy Terms:* The parameters for the consequent dimension ω_{B^i} are calculated according to (4); $\omega_{B^1} = 0.57$, $\omega_{B^2} = 0.43$. The parameters for the missing observation can then be calculated using (21); $\omega_{A_3^1} = 0.65$, $\omega_{A_3^2} = 0.35$, and the intermediate fuzzy set $A_3^\dagger = (5.75, 6.10, 6.45)$ can be obtained with respect to (22). From this, using (5) and (22), the following can be obtained: $\delta_{A_3} = 0.0$, and the shifted fuzzy set $A_3' = (5.75, 6.10, 6.45)$.
2) *Scale and Move Transformation From A_3' to A_3^*:* The individual scale and move parameters can be calculated according to (13) and (14), resulting in $s_B = 1.73$, $m_B = 0.33$. From (24) and (26), it is computed that $s_{A_3} = 1.71$ and $m_{A_3} = 0.75$. The scaled fuzzy term A_3'' is therefore $(5.50, 6.10, 6.70)$. Finally, according to (29), the transformed $A_3^* = (5.65, 5.80, 6.85)$ can be obtained.

Again, the result can be validated by performing the conventional T-FIR using the obtained A_3^*, resulting in the conclusion being $(9.99, 11.00, 12.99)$, which is consistent with the given observed conclusion $(10, 11, 13)$.

Example 3.3: Backward Fuzzy Rule Extrapolation

Extrapolation is a special case of interpolation, when all of the closest rules chosen lie on one side of the hyperplane in which the given observation is a certain point. Determining the closest rules and constructing the intermediate rule are carried out in the same way as those for interpolation. The example below outlines the key steps in the process of backward fuzzy rule extrapolation. Suppose that the observation and the three closest rules as given in Table IV and Fig. 6 are used for extrapolation, where all rules lie on the right side of the observation. In this example, A_2^* is the missing antecedent that is to be extrapolated.

1) *Construction of the Intermediate Fuzzy Terms:* The normalized weights associated with the observed antecedents and conclusion are listed in Table V. The parameters for the missing observation $\omega_{A_2^i}$, $i = 1, 2, 3$ can then be calculated using (21) such that $\omega_{A_2^1} = 0.56$, $\omega_{A_2^2} = 0.27$, $\omega_{A_2^3} = 0.17$, and the intermediate fuzzy set $A_2^\dagger = (5.82, 6.82, 7.82, 8.82)$ can be obtained according to (22). Then, the bias δ_{A_2} between A_2^* and $A_2^\dagger$ is calculated by (22), $\delta_{A_2} = -0.21$. The shifted fuzzy term A_2' which has the same representative value as A_2^*, can be obtained from (5): $A_2' = (2.73, 3.73, 4.73, 5.73)$.
2) *Scale and Move Transformation From A_3' to A_3^*:* The individual scale and move parameters are calculated with respect to (8)–(15), resulting in $\underline{s}_B = 0.64$, $\overline{s}_B = 0.59$, $m_B = -0.61$. The scale ratio $\mathbb{S}_{B^*} = 0.09$ is obtained according to (27). Similarly, the relevant parameters $\underline{s}_{A_k}, \overline{s}_{A_k}, m_{A_k}$ of antecedents A_1^*, A_3^* can be obtained. In particular, using equations similar to (24)–(27), it follows that $\underline{s}_{A_2} = 0.40$, $\overline{s}_{A_2} = 0.50$, $m_{A_2} = -0.91$, and $\mathbb{S}_{A_2} = -0.04$. The scaled fuzzy term A_2'' can then be computed: $(3.27, 3.94, 4.52, 5.19)$. Finally, the required $A_2^* = (3.14, 4.21, 4.79, 5.05)$ is obtained by performing the transformation.

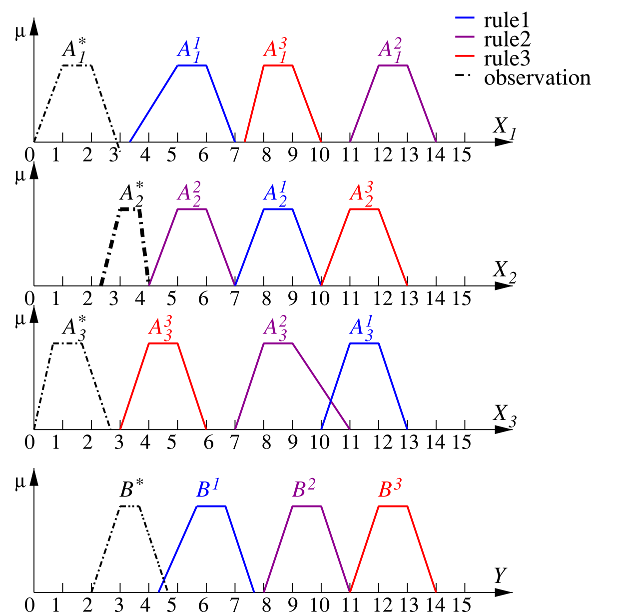


Fig. 6. Example of backward fuzzy rule extrapolation with multiple multi-antecedent rules.

TABLE V
NORMALIZED WEIGHTS OF THE KNOWN ANTECEDENTS

	R_1	R_2	R_3
B	0.55	0.27	0.18
A_1	0.53	0.28	0.19
A_3	0.57	0.25	0.18

Note that as with the previous illustrative cases for interpolation, the above-extrapolated result can be verified to match well with the observation.

IV. BACKWARD FUZZY RULE INTERPOLATION AND EXTRAPOLATION WITH MULTIPLE MISSING ANTECEDENTS

For practical applications, there are often more than one antecedent with missing values. Therefore, the question about how to perform BFRIE with multiple missing values is raised. This section presents two approaches that attempt to address this issue: 1) the parametric approach (see Section IV-A), which directly extends the S-BFRIE method but involves a higher computational complexity, and 2) the feedback approach (see Section IV-B), which is a more generalized method that works more closely with conventional FRI procedures.

A. Parametric Approach

1) Problem Analysis: The key to solving general BFRIE problems, following the principles of the S-BFRIE method, lies with the calculation of the best T-FIR parameter combination, which leads to the closest resemblance of the original (missing) values. In particular, to create the intermediate fuzzy terms that are based on the N closest rules, the following set of parameters:

$$\{(\omega_{A_l^i}, i = 1, 2 \ldots, N), \delta_{A_l}, \underline{s}_{A_l}, \overline{s}_{A_l}, m_{A_l}\} \quad (30)$$

of cardinality $N + 4$ is required to backward interpolate each missing antecedent A_l^*, given trapezoidal representation. Here, $l \in L$, $L \subseteq \{1, \ldots, M\}$, denote the indices of the missing antecedents A_l^*. Taking parameter for bias, δ_{A_l} as an example, the following constraint needs to be satisfied:

$$\sum_{l \in L} \delta_{A_l} = M\delta_B - \sum_{k=1, k \notin L}^{M} \delta_{A_k} \quad (31)$$

which is an extended form of (22) that is used in S-BFRIE. Similar formulae for the remaining parameters may also be derived in the exact same fashion, altogether forming multiple simultaneous equations to be resolved.

Note that apart from those determined by the aforementioned equations, by definition, these parameters also take values from their underlying ranges: $\omega \in [0, 1]$, $\delta \in [-1, 1]$, $\underline{s}_{A_l} \in [0, \infty)$, $\overline{s}_{A_l} \in [0, \infty)$, and $m \in [-1, 1]$. Thus, these ranges need to be discretized in order to generate the required parameter combinations. In assessing the performance of the estimation, the conventional T-FIR procedure can then be invoked to verify the correctness and accuracy. This is done by comparing the output (using the estimated parameters) with the actual observed consequent value so that the most suitable setting may be identified.

After all of the parameters are acquired, in theory, the missing antecedents may be expected to be individually derived using the previously described S-BFRIE steps. Unfortunately, the set of simultaneous equations cannot be resolved in a straightforward manner, due to the lack of sufficient given values. This is because different observations may potentially lead to the same (or very similar) consequent, for any system of a fair complexity. As the number of missing antecedent values increases, the possible scenarios may become extremely wide-reaching or even countless.

From a theoretical point of view, the complexity of this approach mainly comes from the high number of possible parameter combinations (ω, δ, $\underline{s}$, $\overline{s}$, and m), all nonindependent. The weight ω, in particular, is calculated with regard to all $|L|$ missing antecedents and all N closest rules, thus having a considerably high complexity $O(\upsilon^{N|L|})$. Here, $\upsilon \in \mathbb{N}^+, \upsilon > 1$, signifies the number of discretized intervals that are used to generate the possible parameter combinations. Higher υ produces finer intervals and allows closer estimations to the actual values. The discretizations of the other four parameters: δ, $\underline{s}$, $\overline{s}$, and m all have the same computational complexity of $O(\upsilon^{|L|})$. Therefore, the overall computational complexity of generating these combinations is $O(\upsilon^{(N+4)|L|})$, which is prohibitive for large υ, N, and $|L|$. The situation worsens if the verification of every possible parameter combination is required in order to validate that the estimated transformations indeed produce reasonable results, and to enable the selection of better (closer) outputs. The resultant P-BFRIE process will have an overall cost of $O(\upsilon^{(N+4)|L|}) \cdot O(\text{FRI})$, where $O(\text{FRI})$ stands for the complexity of the FRIprocess itself.

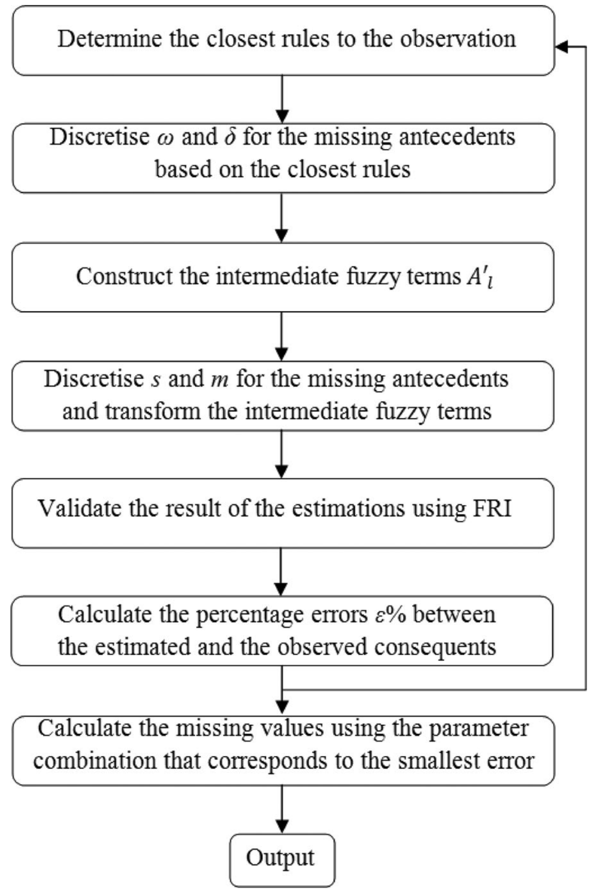


Fig. 7. Flowchart of the parametric approach.

2) Simplified P-BFRIE for Two Missing Antecedent Values: Having undertaken the aforementioned analysis of the computational complexity that theoretical T-FIR involves, practically simpler methods are necessary. For this, a simplified process that supports the interpolation of two missing antecedent values is proposed here, as outlined in Fig. 7. The cost of discretizing ω is reduced to $O(\upsilon^N)$, with an overall time complexity of $O(\upsilon^{N+4}) \cdot O(\text{FRI})$. This simplification takes advantage of the codependence of the two weights $\omega_{A_{l_1}^i}$ and $\omega_{A_{l_2}^i}$, associated with each rule R^i

$$\omega_{A_{l_1}^i} + \omega_{A_{l_2}^i} = M\omega_{B^i} - \sum_{k=1, k \notin L}^{M} \omega_{A_k^i} \quad (32)$$

TABLE VI
TWO CLOSEST RULES FOR OBSERVATION

	O	R_1	R_2
x_1	(2.88, 3.50,3.90, 4.66)	(3.07,3.55,3.73,4.08)	(3.51,3.75,3.99,4.23)
x_2	(0.96, 1.29,1.56, 1.97)	(0.88,1.21,1.32,1.47)	(1.57,2.19,2.29,3.01)
x_3	missing	(1.23,1.64,2.08,2.65)	(1.41,1.83,2.22,2.68)
x_4	(8.37,8.56, 9.24, 9.68)	(7.91,8.34,8.87,9.79)	(8.03,8.62,8.90,9.14)
x_5	missing	(0.39,0.88,1.29,1.97)	(0.94,1.41,1.92,2.34)
y	(9.56,10.54,10.83,11.58)	(11.36,12.40,12.66,12.70)	(12.77,13.83,14.54,15.19)

TABLE VII
PARAMETERS FOR ESTIMATING THE MISSING ANTECEDENTS

A_l	$\omega_{A_l^1}$	$\omega_{A_l^2}$	δ_{A_l}	$\underline{s}_{A_l}$	$\overline{s}_{A_l}$	m_{A_l}
$l = 3$	0.92	0.08	0.82	0.61	0.90	-1.0
$l = 5$	0.39	0.61	-0.27	1.13	0.15	0.23

where one value uniquely determines the other during the discretization process.

The outcomes of the now more manageable υ^{N+4} parameter combinations can then be verified through FRI, where a simple measurement of percentage error $\epsilon^j_{\%}$ may be used

$$\epsilon^j_{\%} = 100 \times d(B^{j*}, B^*). \tag{33}$$

Here, the two estimated missing antecedent values $A^{j*}_{l_1}$ and $A^{j*}_{l_2}$, backward interpolated through the use of the jth parameter combination, are employed to obtain a certain B^{j*}. The distance between this estimated consequent and the actual observed consequent B^* determines the accuracy of the transformations that have taken place. Finally, the set of A^{j*}_l, which corresponds to the minimal resulting error, is chosen as the desired output, since it is the best approximation possible given the limited information, and the amount of discretized intervals employed.

Example 4.1: Table VI lists the observation and the closest rules chosen according to (20), where x_3 and x_5 are assumed to have two missing antecedent values. For this particular example, if $\upsilon = 12$ is used to discretize each parameter, the total number of parameter combinations ω, δ, s, and m is $12^{(2+4)}$ (with $N = 2$), resulting in the same number of pairs of possible A^{j*}_3 and A^{j*}_5. The parameters corresponding to the validated consequent with the smallest error $\epsilon^j_{\%} = 0.22\%$ [according to (33)] are listed in Table VII, where $B^{j*} = (12.70, 13.03, 13.63, 14.66)$. The final backward interpolative values for the two missing antecedents are $A^*_3 = (2.02, 2.18, 2.40, 2.77)$ and $A^*_5 = (0.52, 2.17, 3.03, 3.55)$.

B. Feedback Approach

This section describes an alternative and more intuitive approach to BFRIE, termed the feedback approach (F-BFRIE). It significantly reduces the time-complexity for parameter estimation. This is shown in the flowchart of Fig. 8 and illustrated in Algorithm 1. It works by directly estimating the possible initial values of the missing antecedents, then validating the resultant consequent through conventional FRI, in order to identify the most suitable value combination(s) that lead to the observed consequent value. For consistency and ease of explanation, as

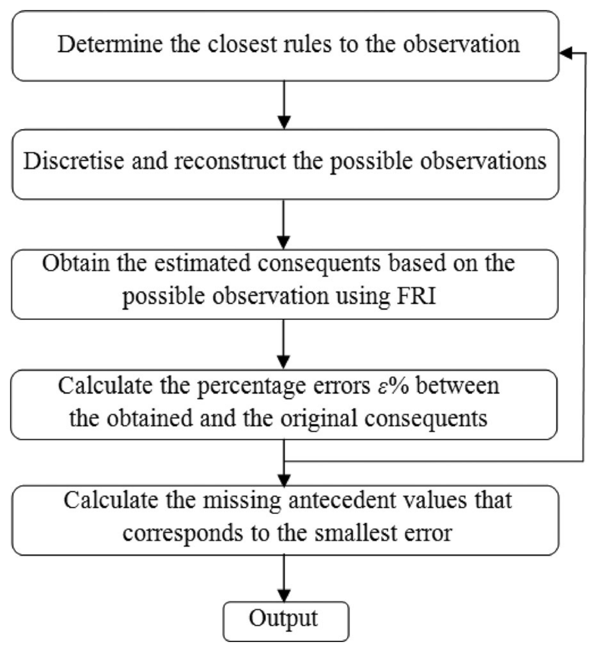


Fig. 8. Flowchart of the feedback approach.

Algorithm 1: Feedback Approach

for $l \in L$ **do**
  $o = \emptyset$
  for $c^j_l = \min_{A_l}$ **to** $\max_{A_l}$ **do**
    $A^{j*}_l = approx(c^j_l)$
    $o = o \cup A^i_l$
    $c^j_l = c^j_l + range_{A_l}/\upsilon$
  combinations = combinations $\cup\, o$
return *combinations*

with P-BFRIE, mechanisms such as T-FIR, discretization of variable ranges, and the percentage error-based validation are again used in the implementation.

In order to obtain the initial estimation, the domain ranges of the missing antecedents themselves (rather than those of the parameters previously used for P-BFRIE) are discretized into υ intervals. The resulting crisp points are then used to approximate a total of $\upsilon^{|L|}$ possible value combinations for the missing antecedent variables $\{A^{j*}_l\}, l \in L, j = 1, \ldots, \upsilon^{|L|}$. Assume a given crisp point c^j_l for the lth antecedent variable, (34)–(36) detail this approximation procedure, which is denoted by approx(c^j_l) in Algorithm 1

TABLE VIII
CLOSEST RULES FOR TWO EXAMPLE RECONSTRUCTED OBSERVATIONS O^{p*} AND O^{q*}, $p, q \in 1, \ldots, v^{|L|}$

	O^{p*}	R_1	R_2	R_3
x_1	$A_1^{p*} = (6.12, 6.70, 7.28, 7.55)$	(7.68,8.18,9.18,9.68)	(8.06,8.56,9.56,10.06)	(7.84,8.34,9.34,9.84)
x_2	(1.32,1.74,2.20,2.78)	(6.09,6.59,7.59,8.09)	(0.64,1.14,2.14,2.64)	(8.68,9.18,10.18,10.68)
x_3	$A_3^{p*} = (7.37, 7.95, 8.53, 9.03)$	(7.91,8.41,9.41,9.91)	(6.47,6.97,7.97,8.47)	(0.39,0.89,1.89,2.39)
x_4	(4.36,4.58,4.63,5.32)	(4.75,4.89,5.12,5.73)	(5.11,5.14,5.24,5.89)	(3.51,4.38,4.58,5.00)
x_5	$A_5^{p*} = (7.72, 8.29, 8.53, 8.70)$	(5.86,6.36,7.36,7.86)	(0.15,0.65,1.65,2.15)	(6.72,7.22,8.22,8.72)
x_6	(4.19,4.47,5.35,5.80)	(7.76,8.26,9.26,9.76)	(8.08,8.58,9.58,10.08)	(0.60,1.10,2.10,2.60)
x_7	$A_7^{p*} = (3.64, 4.22, 4.80, 5.37)$	(5.92,6.42,7.42,7.92)	(0.67,1.17,2.17,2.67)	(1.52,2.02,3.02,3.52)
y		(5.79,6.29,7.29,7.79)	(7.73,8.23,9.23,9.73)	(0.31,0.81,1.81,2.31)
	O^{q*}	R_1	R_2	R_3
x_1	$A_1^{q*} = (1.73, 2.70, 3.57, 3.97)$	(1.00,1.55,2.27,2.78)	(5.03,5.73,6.26,7.01)	(8.29,8.33,8.70,9.09)
x_2	(1.32,1.74,2.20,2.78)	(1.44,1.87,2.38,3.00)	(3.74,4.24,4.56,5.02)	(3.50,4.37,4.83,5.25)
x_3	$A_3^{q*} = (2.33, 3.53, 4.51, 5.58)$	(8.72,9.06,9.65,10.04)	(7.57,7.99,8.46,8.84)	(7.44,7.98,8.58,8.90)
x_4	(4.36,4.58,4.63,5.32)	(4.75,4.89,5.12,5.73)	(5.11,5.14,5.24,5.89)	(3.51,4.38,4.58,5.00)
x_5	$A_5^{q*} = (4.73, 5.71, 6.61, 7.44)$	(6.47,6.89,7.53,8.01)	(4.17,4.94,5.67,6.32)	(5.57,5.95,6.38,7.17)
x_6	(4.19,4.47,5.35,5.80)	(6.47,6.89,7.53,8.08)	(5.77,5.86,6.34,7.24)	(5.32,5.84,6.74,7.77)
x_7	$A_7^{q*} = (3.87, 4.65, 5.23, 5.79)$	(1.78,2.68,3.58,4.12)	(4.91,5.51,6.31,6.87)	(6.01,6.07,6.26,7.33)
y		(8.00,9.04,9.64,10.08)	(6.09,6.67,7.19,7.34)	(4.79,5.57,6.31,6.62)

$$\begin{cases} a_{0_{A_l^{j*}}} = c_l^j - \dfrac{\sum_{i=1}^{N}(\mathrm{Rep}(A_l^i) - a_{0\,A_l^i})}{N} \cdot \underline{\Delta} \\ a_{1_{A_l^{j*}}} = c_l^j - \dfrac{\sum_{i=1}^{N}(\mathrm{Rep}(A_l^i) - a_{1\,A_l^i})}{N} \cdot \overline{\Delta} \\ a_{2_{A_l^{j*}}} = c_l^j + \dfrac{\sum_{i=1}^{N}(a_{2\,A_l^i} - \mathrm{Rep}(A_l^i))}{N} \cdot \overline{\Delta} \\ a_{3_{A_l^{j*}}} = c_l^j + \dfrac{\sum_{i=1}^{N}(a_{3\,A_l^i} - \mathrm{Rep}(A_l^i))}{N} \cdot \underline{\Delta} \end{cases} \tag{34}$$

where

$$\underline{\Delta} = \frac{N\left(\sum_{k=1,k\notin L}^{M} \frac{a_{3\,A_k^*} - a_{0\,A_k^*}}{\sum_{i=1}^{N}\left(a_{3\,A_k^i} - a_{0\,A_k^i}\right)} + \frac{a_{3\,B^*} - a_{0\,B^*}}{\sum_{i=1}^{N}\left(a_{3\,B^i} - a_{0\,B^i}\right)}\right)}{M - |L| + 1} \tag{35}$$

$$\overline{\Delta} = \frac{N\left(\sum_{k=1,k\notin L}^{M} \frac{a_{2\,A_k^*} - a_{1\,A_k^*}}{\sum_{i=1}^{N}\left(a_{2\,A_k^i} - a_{1\,A_k^i}\right)} + \frac{a_{2\,B^*} - a_{1\,B^*}}{\sum_{i=1}^{N}\left(a_{2\,B^i} - a_{1\,B^i}\right)}\right)}{M - |L| + 1} \tag{36}$$

where M is the total number of antecedent variables and N is the number of the closest rules.

To obtain close approximations of the missing values, it is useful to remember that the estimated fuzzy sets A_l^{j*} are influenced by both the selected rules: $R_i, i = 1, \ldots, N$, and the observed values $A_k^*, k \notin L$ and B^*. In particular, for a trapezoidal fuzzy set A_l^{j*} that may be returned as the estimated outcome, the positions of its four points are defined relative to the averaged (over the N closest rules) displacements between their corresponding points and the representative values of the fuzzy antecedents $A_l^i, i = 1, \ldots, N$. The points $a_{0A_l^{j*}}$ and $a_{3A_l^{j*}}$ are then scaled with respect to the ratio $\underline{\Delta}$, which is calculated from the averaged (over the N closest rules and all known antecedent/consequent dimensions) ratios between the supports of the observed values, and those of the existing rules. Similarly, $a_{1A_l^{j*}}$ and $a_{3A_l^{j*}}$ are adjusted with respect to $\overline{\Delta}$.

TABLE IX
ERRORS $\epsilon_{\%}$ BETWEEN B^{p*}, B^{q*}, AND B^*

Consequent	Value	$\epsilon_{\%}$
B^*	(8.46,9.26,9.53,9.83)	
⋮	⋮	⋮
B^{p*}	(9.74,10.16,10.93,11.08)	6.14%
⋮	⋮	⋮
B^{q*}	(8.09,8.54,9.60,10.10)	0.74%
⋮	⋮	⋮

The $v^{|L|}$ possible combinations of the fuzzy sets being estimated are used to obtain their respective consequent values $B^{j*}, j = 1, \ldots, v^{|L|}$, through the conventional T-FIR procedure. Note that the closest rules chosen for each of the combinations may be different, since the distance calculation is purely based on the values of the currently estimated observations. The percentage error $\epsilon_{\%}^j$ is then calculated using (33), and the estimated missing antecedent values corresponding to the smallest $\epsilon_{\%}^j$ are returned as the final result.

The computational complexity of this approach is principally due to the generation process of the initial fuzzy sets $O(v^{|L|})$, which is much more scalable than that of P-BFRIE: $O(v^{(N+4)|L|})$. The run-time cost of F-BFRIE is also independent of the number of closest rules N, and of course, the overhead incurred by estimating the T-FIR parameters as required in P-BFRIE is also eliminated.

Example 4.2: Consider the observation givens in Table VIII, where the values of x_1, x_3, x_5, and x_7 are assumed to be missing. According to (34), a total of 20^4 ($v = 20, |L| = 4$) possible value combinations are used to generate the same number of potential consequent values. O^{p*} and O^{q*}, which are two of such combinations obtained in the process and the different closest rules chosen using (20) are shown in Table VIII. After forward interpolation with these approximated fuzzy

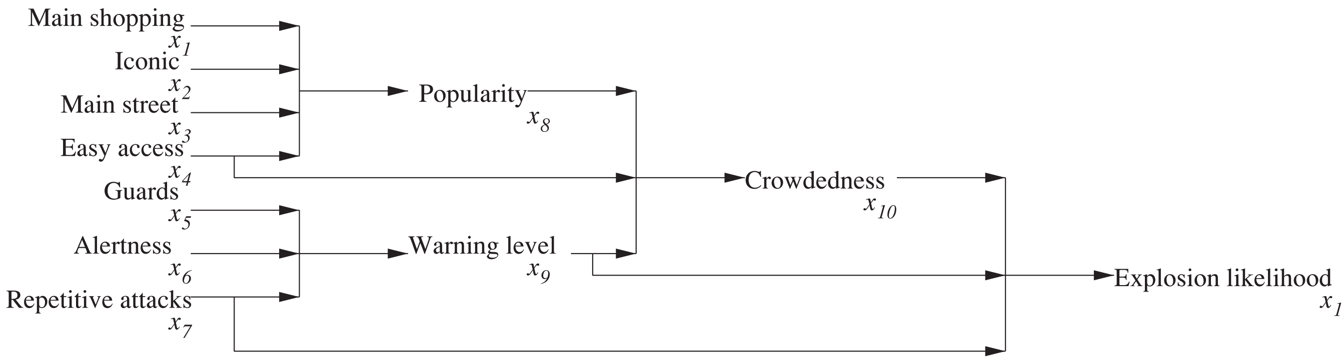


Fig. 9. Causal network model.

sets, the $\epsilon_{\%}^{j}$ obtained during this example evaluation is given in Table IX. The smallest $\epsilon_{\%}$ is 0.74%, and the corresponding consequent $B^{q*} = (9.74, 10.16, 10.93, 11.08)$. The resultant estimated outcome of the four missing antecedents $A_1^* = (1.73, 2.70, 3.57, 3.97)$, $A_3^* = (2.33, 3.53, 4.51, 5.58)$, $A_5^* = (4.73, 5.71, 6.61, 7.44)$, and $A_7^* = (3.87, 4.65, 5.23, 5.79)$ is therefore the final BFRIE result.

V. Experimentation and Discussion

In this section, a practical problem concerning the prediction of terrorist bombing threats is employed to demonstrate the potential of the proposed study. It shows how the implemented techniques help to interpolate the final conclusion, when the system is presented with partial observations, including the reflection of interesting characteristics of the proposed study. In addition, a comparative study of the parametric and feedback methods is also included, which is systematically conducted using a randomized numerical problem, in terms of their approximation accuracy and run-time efficiency.

A. Prediction of Terrorist Bombing Threats

1) Problem Specification and Model Construction: To solve the puzzle of a serious crime including terrorist attacks from a set of given evidence, investigators aim to reconstruct the possible scenarios that may have taken place. The bottleneck of accomplishing this task is the fact that humans are relatively inefficient at hypothetical reasoning, especially when a hierarchically structured procedure is involved. A fuzzy decision-support system may assist investigators in generating plausible scenarios and analyzing them objectively [15], [41]. In this paper, a real-world scenario that involves the prediction of a terrorist threat is considered, which is based on a recent study on suicide bombings [40]. The causal network model used in this experiment is illustrated in Fig. 9. There are 11 variables in this problem, denoted as $x_p, p = 1, 2, \ldots, 11$, involved in four subsets of rules.

1) The first subset concerns the *Popularity* of a place. *Main shopping* (x_1) and *Iconic* (x_2) indicate whether the location is a principal shopping area, or a site of symbolic value, respectively. *Main street* (x_3) describes how busy the area may become. The variable *Easy access* (x_4) refers to the convenience of transportation. These factors jointly determine the *Popularity* (x_8) of a given location.

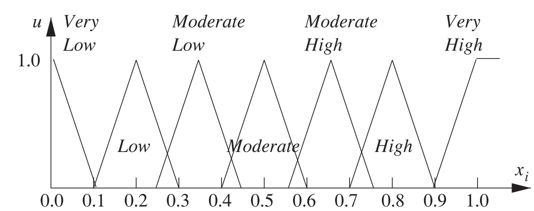


Fig. 10. Definition of the linguistic terms for domain variables.

2) The second subset of rules deals with the *Warning level* (x_9), or the amount of risk to the attackers, which is related to the number of security *Guards* (x_5) in the area, the *Alertness* (x_6) of people, and the numbers of *Repetitive attacks* (x_7) in the past.
3) The third focuses on the prediction of *Crowdedness* (x_{10}). The number of people in an area is directly related to the *Popularity* of the place, and can also be affected by the level of *Easy access*. In addition, the *Crowdedness* may change in relation to the *Warning level*, since cautious individuals may shy away from places that are considered dangerous.
4) The fourth and final rule subset is about the *Explosion likelihood* (x_{11}), which is in indirect relation to the number of people in the area. In addition, the amount of public warning signs displayed in the area may discourage potential attacks, as people are more alert to the surroundings, and suspicious individuals or items may be promptly reported. Moreover, terrorists typically target certain places that repetitively draw their attention, instead of any locations at random.

A selection of the original rules contained in the rule base are given in Table X. In this model, fuzziness is naturally obtained from the presence of the linguistic terms that describe the real-valued domain variables. For simplicity, triangular fuzzy membership functions are applied in this scenario. Note that different variables are defined on their own underlying domains. To simplify knowledge representation, these domains are normalized in this experiment with a range of 0 to 1. The fuzzy sets that represent the normalized linguistic terms are given in Fig. 10. It is important to note that the original rule base consists of substantial gaps, which makes interpolation essential.

2) Work Flow of BFRIE in Action: The set of observations used in this experiment is given in Table XI, where the values

TABLE X
EXAMPLE RULES

	Popularity				
	Main shopping (x_1)	*Iconic* (x_2)	*Main street* (x_3)	*Easy access* (x_4)	*Popularity* (x_8)
R_1	VL	L	L	L	L
R_2	L	H	L	M	ML
R_3	L	M	M	H	M
R_4	M	M	M	L	ML
R_5	M	H	M	MH	M
R_6	M	M	M	H	MH
R_7	H	L	L	L	M
R_8	H	M	L	M	MH
R_9	H	M	M	H	H
R_{10}	VH	H	H	M	H

	Warning level			
	Guards (x_5)	*Alertness* (x_6)	*Repetitive attacks* (x_7)	*Warning level* (x_9)
R_1	L	VL	VL	VL
R_2	L	M	L	L
R_3	L	H	M	M
R_4	ML	M	L	L
R_5	M	L	M	ML
R_6	M	L	L	ML
R_7	M	M	M	M
R_8	M	H	M	M
R_9	H	M	M	MH
R_{10}	H	H	H	H

	Crowdedness			
	Popularity (x_4)	*Easy access* (x_8)	*Warning* (x_9)	*Crowdedness* (x_{10})
R_1	VL	VL	VH	VL
R_2	H	VH	VL	VH
R_3	L	H	H	L
R_4	H	M	M	MH
R_5	M	M	H	L
R_6	ML	H	L	M
R_7	M	H	H	M
R_8	H	VL	H	ML
R_9	MH	H	M	MH
R_{10}	H	H	VL	VH

	Explosion likelihood			
	Crowdedness (x_7)	*Warning level* (x_9)	*Repetitive attacks* (x_{10})	*Explosion likelihood* (x_{11})
R_1	VL	L	L	VL
R_2	L	M	M	L
R_3	L	L	H	ML
R_4	M	L	M	MH
R_5	M	M	H	M
R_6	M	H	M	L
R_7	H	L	M	MH
R_8	H	L	H	VH
R_9	H	M	M	MH
R_{10}	VH	M	L	H

VL: Very Low, L: Low, ML: Moderate Low, M: Moderate, MH: Moderate High, H: High, VH: Very High.

TABLE XI
OBSERVATIONS

O_1	x_1	x_2	x_3	x_4	x_8
	H	H	MH	missing	H
O_2	x_5	x_6	x_7	x_9	
	missing	L	MH	missing	
O_3	x_4	x_8	x_9	x_{10}	
	missing	H	missing	M	
O_4	x_7	x_9	x_{10}	x_{11}	
	ML	missing	M	?	

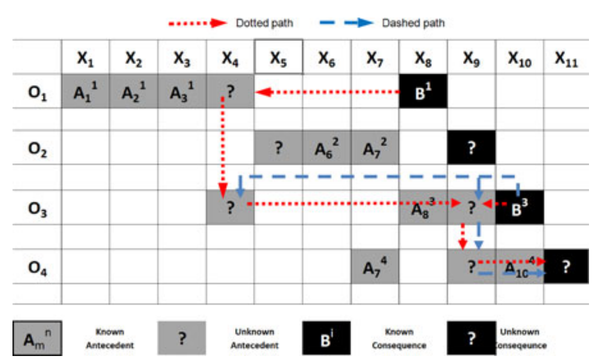


Fig. 11. Example structure for bombing attack prediction.

of the antecedent variables *Easy access* (x_4), *Guards* (x_5), and *Warning level* (x_9) are all missing, as well as that of the final consequent x_{11} to be interpolated. The antecedent variable *Warning level* (x_9) is of particular importance, since it is involved in two subsets of rules (for *Crowdedness* and *Explosion likelihood*). Without x_9, no matter what other information is available, even with the *Repetitive attack* (x_7) and *Crowdedness* (x_{10}) known, forward interpolation will still fail.

It is not uncommon for a hierarchical reasoning framework to have more than one path of inference/interpolation [17], [36]. For this particular set of observations, it is possible to obtain the value of x_9 through two different paths, as shown in Fig. 11.

1) *Dotted Path via S-BFRIE*:
   a) Calculate the value of *Easy access* (x_4) according to the given consequence value x_8 and the antecedent values x_1, x_2, and x_3, using the subrule base *Popularity* via S-BFRIE. Following the steps detailed previously, the backward interpolated value is $x_4 = (0.42, 0.52, 0.62)$ (*M*).
   b) Interpolate the value of x_9 using the subrule base *Crowdedness*. The values of x_8 and x_{10} are directly observed, and that of x_4 is obtained from the previous step. The three closest rules are then selected using the consequence-biased distance measure as per (20). The resulting closest rules with respect to the observation, and the backward interpolated value $x_9 = (0.52, 0.62, 0.72)$ (*MH*), are shown in Fig. 12.
   c) Use the interpolated value of x_9, and the other given values for x_7 and x_{10} to forward interpolate the final consequent variable *Explosion likelihood*

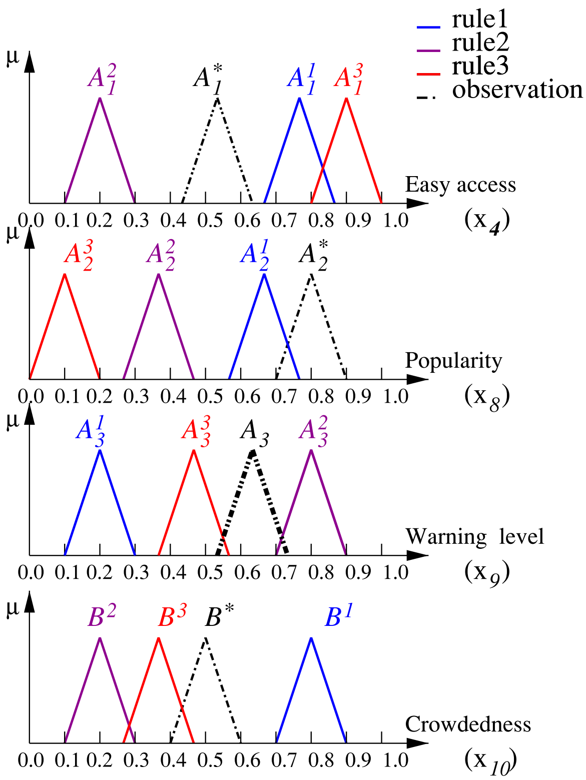


Fig. 12. Calculate the *Warning level* using S-BFRIE.

(x_{11}). This leads to the required final result $x_{11} = (0.61, 0.71, 0.81)$ (*MH*).

2) *Dashed Path via P-BFRIE and F-BFRIE*

a) Calculate x_4 and x_9 simultaneously, according to the given values x_8 and x_{10}, and the subrule base *Crowdedness*. By employing the P-BFRIE approach, the results obtained are $x_4 = (0.46, 0.56, 0.66)$ (*M*) and $x_9 = (0.51, 0.61, 0.71)$ (*MH*). Alternatively, the results obtained using F-BFRIE are $x_4 = (0.43, 0.53, 0.63)$ (*M*) and $x_9 = (0.59, 0.69, 0.79)$ (*MH*). Both agree with the previous results via the use of S-BFRIE: $x_4 = (0.42, 0.52, 0.62)$ (*M*) and $x_9 = (0.52, 0.62, 0.72)$ (*MH*).

b) Derive, in a manner similar to the dotted path above, the value of the final consequent variable $x_{11} = (0.58, 0.68, 0.78)$ (*MH*) (according to P-BFRIE: $x_9 = (0.51, 0.61, 0.71)$) and $x_{11} = (0.56, 0.66, 0.76)$ (*MH*) (according to F-BFRIE: $x_9 = (0.59, 0.69, 0.79)$). The errors $\epsilon_{\%}$ between the result of the dotted path ($x_{11} = (0.61, 0.71, 0.81)$) and these are 3.0% and 5.0%, respectively. This demonstrates that both processing paths are feasible for dealing with this problem.

3) Note that after the aforementioned interpolation process, the last remaining missing value for *Guards* (x_5), although no use in this particular prediction application, can also be backward interpolated. In particular, by using the observed values of x_6 and x_7, and the interpolated value of x_9, the result of $x_5 = (0.28, 0.38, 0.48)$ (*ML*) can be obtained via the use of subrule base *Warning level*.

TABLE XII
OBSERVATION USED FOR THE INVESTIGATION OF MULTIPLE POSSIBLE OUTCOMES

Main shopping (x_1)	*Iconic* (x_2)	*Main street* (x_3)	*Easy access* (x_4)	*Popularity* (x_8)
missing	missing	M	ML	MH

3) Practical Significance of BFRIE: In real applications, it is often difficult to predict and adjust the *Warning level* until (suicide) bombing attacks have actually occurred or prevented. However, with BFRIE, the *Warning level* may now be estimated from the other related factors. This may significantly increase the effectiveness of the predication and prevention of bombing attacks. However, *Easy access* and *Guards* are controllable elements which may be adjusted in order to minimize the *Explosion likelihood*.

In order to reduce the *Explosion likelihood* (x_{11}) from $(0.61, 0.71, 0.81)$ (*MH*) say, to $(0.30, 0.40, 0.50)$ (*ML*), according to the proposed P-BFRIE method, the value of *Easy access* (x_4) needs to be changed from the current $(0.42, 0.52, 0.62)$ (*M*) to $(0.23, 0.33, 0.43)$ (*ML*), and similarly, *Warning level* (x_9) from the current $(0.52, 0.62, 0.72)$ (*MH*) to $(0.74, 0.83, 0.94)$ (*H*), and *Guards* (x_5) from the current $(0.28, 0.38, 0.48)$ (*ML*) to $(0.54, 0.64, 0.74)$ (*MH*). Thus, with the use of the proposed reverse reasoning technique to interpolate the crucial variables, the risk of a certain area concerned may be significantly reduced (or future repetitive attacks prevented).

4) Use of Alternative Distance Metrics: If the unbiased distance measure (3) is used to backward interpolate the missing value x_9, the same closest rules will no longer be selected. Instead a different outcome of *Warning level* $x_9 = (0.18, 0.28, 0.38)$, and *Explosion likelihood* $x_{11} = (0.29, 0.39, 0.49)$ (*ML*) will be returned. Looking back at the original observation, given the two antecedent values such that *Easy access* (x_4) is *M* and *Popularity* (x_8) is *H*, the intuitive deduction of *Crowdedness* (x_{10}) should be quite high, as the place is both moderately high in popularity and reasonably convenient to reach. The only reason why the observed *Crowdedness* (x_{10}) has a *moderate* value may well be because of a reasonable *Warning level*. Therefore, the outcome $x_9 = (0.52, 0.62, 0.72)$ from the use of the biased distance measure is more agreeable than $x_9 = (0.18, 0.28, 0.38)$ resulting from the use of the plain distance measure. Experiments show that, if $x_9 = (0.18, 0.28, 0.38)$ and the antecedent values of x_4 and x_8 are *M* and *H* respectively, T-FIR method will result in an interpolated $x_{10} = (0.14, 0.24, 0.34)$ (*L*), which will be much further than the original observation: *Crowdedness* (x_{10}) is *M*. This clearly demonstrates the significance in utilizing the biased distance metric proposed in this paper.

5) Multiple Equally Probable Interpolative Outcomes: The involvement of multiple missing antecedents naturally implies that alternative equally probable combinations of observations may be present, which may all lead to the same consequent observed. Assume that the observation shown in Table XII is

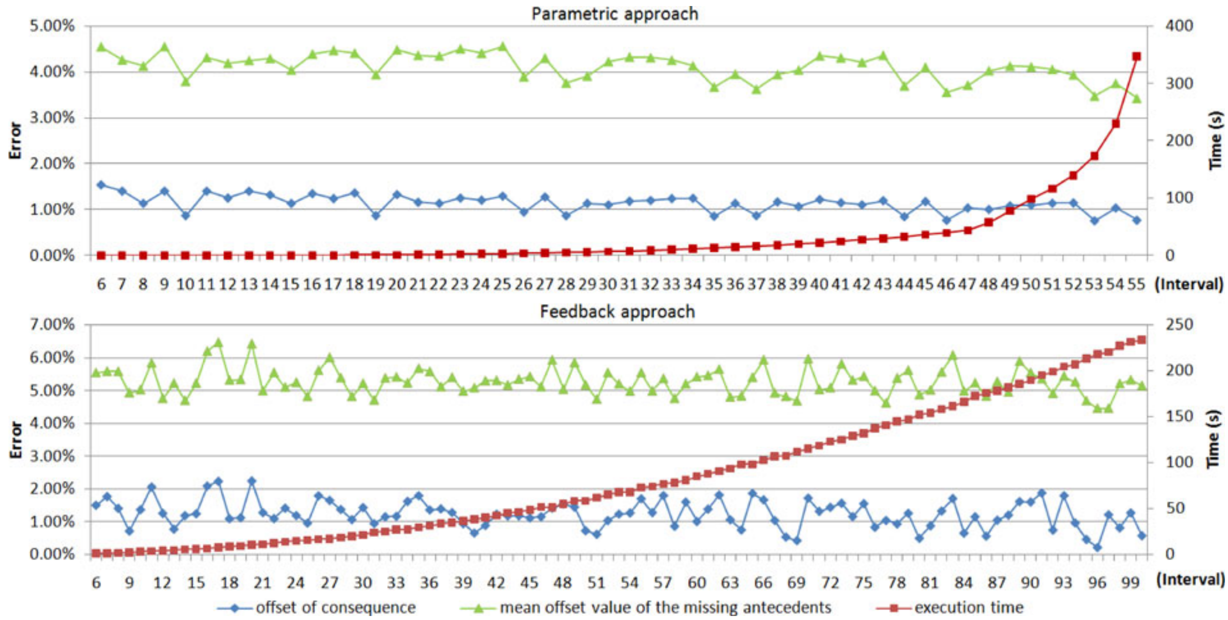


Fig. 13. Relationship between the value of υ, approximation error $\epsilon_{\%}$, and execution time.

TABLE XIII
RELATIONSHIP BETWEEN *MAIN SHOPPING* AND *ICONIC*

$\epsilon_{\%}$	*Main shopping*		*Iconic*	
0.29	(0.66,0.76,0.86)	H	(0.50,0.60,0.70)	MH
0.56	(0.54,0.54,0.64)	M	(0.72,0.82,0.92)	H
0.91	(0.19,0.29,0.39)	VL	(0.76,0.86,0.96)	H
1.06	(0.50,0.60,0.70)	MH	(0.74,0.84,0.94)	H
1.20	(0.67,0.77,0.87)	H	(0.70,0.80,0.90)	H
1.28	(0.70,0.80,0.90)	H	(0.20,0.30,0.40)	ML

TABLE XIV
ERRORS OF P-BFRIE AND F-BFRIE OVER 500 TEST SAMPLES

	P-BFRIE			F-BFRIE		
$\epsilon_{\%}$	B	A_3	A_5	B	A_3	A_5
Max	11.32	9.57	8.39	12.70	10.30	10.33
Min	0.0000	0.0073	0.0090	0.0005	0.0004	0.0039
Mean	1.68	3.67	3.12	1.63	5.72	4.95

given, where the values of *Main shopping* and *Iconic* are both missing. The P-BFRIE is adopted to calculate these two missing antecedent values and the different results found with similar errors ($\epsilon_{\%} < 1.50\%$) are summarized in Table XIII. These results reveal that the same value of *Popularity* can be obtained when either *Main shopping* or *Iconic* is *High* or *Very High*. This also agrees with the intuition, since either of these variables may cause a given location to be attractive, and both may be equally effective to influence the final outcome. Theoretically speaking, this particular subset of rules contains redundant variables, and may be further pruned for higher efficiency. The presence of such redundancy may prompt the use of dimensionality reducing techniques such as feature selection [12], [22].

B. Comparative Studies

To systematically compare the two proposed methods: P-BFRIE and F-BFRIE, a numerical function shown in (37) is used. The rule base employed in the experimental evaluation is generated using the following steps: 1) a random set of crisp values are selected for the function variables and the outcome is calculated according to (37); 2) these crisp values are then fuzzified into trapezoidal fuzzy sets; and 3) the rule base is then populated using these randomly generated rules, while checking (and where appropriate, removing) rules to ensure the underlying domain is reasonably covered, while there still exist sufficient "gaps" between rules in order to utilize interpolation

$$y = 3x_1 - 3.3x_2 + 0.4x_3 + 0.5x_4 + 0.7x_5. \qquad (37)$$

This experimental setup enables an initial sparse rule base to be generated that is an approximation of the underlying knowledge, simulating those obtainable by "subject experts." An observation is obtained in a similar manner, where the "missing" values are then purposefully removed to facilitate backward reasoning. Since the underlying function, i.e., "ground truth," is available. The consistency, accuracy, and robustness of the interpolative procedure can then be verified by comparing the outcome of the interpolation to the actual value computed using (37). This test, therefore, reflects an underlying principle similar to that behind cross validation and statistical evaluation [3], [33].

Altogether, 500 simulated samples are randomly drawn from the domain $U = [0, 10]^5$. Without losing generality, the values of x_3 and x_5 are assumed to be missing. The errors for the consequent and the missing antecedents over these testing records are summarized in Table XIV. For the consequent variable B, the errors are obtained by calculating the distance between the estimated consequent B^{j*} and the actual value B^*. The errors of the two antecedent variables with missing values A_3 and A_5 are derived from the distances between the interpolative outcomes (i.e., values corresponding to the smallest consequent error) and the actual values of A_3^* and A_5^* (the ground truths). It can be seen that the parametric approach demonstrates a higher accuracy than F-BFRIE. This is likely because the parametric approach has precise control of the parameter values. In addition, the shapes of the initial fuzzy sets used in F-BFRIE are approximated, which may have affected its performance. Nevertheless, both methods seem to have an acceptable level of errors.

The value of υ (the number of discretized intervals per variable) is an important factor for both approaches. Fig. 13 presents the relationships between the approximation errors and the execution times with respect to various values of intervals υ, for the two proposed methods. The results show that the parametric approach produces a higher accuracy when a larger number of intervals is used. However, it is also less scalable. This experimentally demonstrates that the run-time complexity and memory requirement of F-BFRIE are more relaxed, which also agrees with the theoretical analysis regarding their complexities in Sections IV-A and IV-B.

VI. Conclusion

This paper has presented BFRIE, a novel approach that complements traditional FRI by supporting backward inference, allowing flexible interpolation when certain antecedents are missing from the observation. This study is based upon the mechanisms of T-FIR, in order to handle multiple multiantecedent rules and to ensure the maintenance of convexity and normality of interpolated outcomes. Specific algorithms have been developed to tackle problem scenarios with single and multiple missing values, with worked examples provided to illustrate their operations. Its practical significance and potential have been demonstrated with a real-world problem scenario: the prediction of terrorist bombing attacks. Systematic evaluation results also show that P-BFRIE is more accurate, despite its limited scalability for larger problems. The F-BFRIE can successfully handle more complex scenarios, and significantly reduces the need of parameter calculation, but its interpolative accuracy is relatively lower.

The current techniques are implemented using exhaustive search-based methods, which may be better formulated with advanced solution techniques (e.g., Waltz algorithm [47]) for flexible constraint satisfaction [37]. It may also be further improved via the use of heuristic optimization algorithms [12], [35], [48] that do not require domain discretization. This may help to obtain even better results, while reducing the search cost (for P-BFRIE in particular). This study will also benefit from a mechanism for automatic identification and selection of better reasoning paths in handling hierarchical rule models so that the interpolation process may dynamically proceed [38] according to the current states of the system.

The underlying concept that supports the proposed BFRIE seems to bear close relation to that of fuzzy inversion [2], [45], [46]. A systematic comparison between these two approaches is, however, beyond the scope of this paper and remains active research. It is also worth extending the BFRIE approach to support other types of interpolation methods (e.g., GM [1], FIVE [31], IMUL [49]). While in principle, the idea of BFRIE (or backward reasoning in general) appears to be applicable to both Mamdani and TSK fuzzy systems [14], [18], [25], [26], for the present implementation, the technique described relies on the scale and move transformation-based procedures and is therefore, only applicable to Mamdani models. It is of natural appeal to develop the proposed technique for TSK fuzzy models. Intuitively, it may also be interesting to apply the technique to different problem domains, such as network intrusion detection [42] and oil exploration [49].

As indicated previously, the underlying problem that BFRIE addresses is that of "many to one," where a number of different value combinations, for the antecedent variables, may lead to very similar observed values for the consequent variable. Although this issue has been partly analyzed from an experimental viewpoint, much remains to be done. In particular, problem may exacerbate if the number of missing values becomes larger. This makes it very challenging to restore the "true" original observation. Fortunately, different observations obtained via the BFRIE process will generate the same or very similar outcomes. Thus, they may be regarded as "equally possible" given the limited amount of knowledge with regard to the application problem. Nevertheless, it is important to be able to improve the proposed approach in an effort to better handle the "many to one" problem. This will be of practical significance for BFRIE to be employed by accuracy-critical applications (e.g., medical diagnosis [16]). The proposed methods may be further extended and combined with the adaptive fuzzy interpolation technique which ensures inference consistency [52]. Furthermore, an intelligent antecedent and/or rule selection procedure may be developed by identifying the most relevant information [5], [11], [12], [39] so that the appropriate terms or rules can be determined to minimize the overall system complexity.

References

[1] P. Baranyi, L. T. Kóczy, and T. D. Gedeon, "A generalized concept for fuzzy rule interpolation," *IEEE Trans. Fuzzy Syst.*, vol. 12, no. 6, pp. 820–837, Dec. 2004.

[2] P. Baranyi, P. Korondi, H. Hashimoto, and M. Wada, "Fuzzy inversion and rule base reduction," in *Proc. Int. Conf. Intell. Eng. Syst.*, 1997, pp. 301–306.

[3] G. Bontempi, H. Bersini, and M. Birattari, "The local paradigm for modeling and control: From neuro-fuzzy to lazy learning," *Fuzzy Sets Syst.*, vol. 121, no. 1, pp. 59–72, 2001.

[4] T. Boongoen, Q. Shen, and C. Price, "Disclosing false identity through hybrid link analysis," *Artif. Intell. Law*, vol. 18, no. 1, pp. 77–102, 2010.

[5] T. Boongoen and Q. Shen, "Nearest-neighbor guided evaluation of data reliability and its applications," *IEEE Trans. Syst., Man, Cybern.*, vol. 40, no. 6, pp. 1622–1633, Dec. 2010.

[6] B. Bouchon-Meunier, R. Mesiar, C. Marsala, and M. Rifqi, "Compositional rule of inference as an analogical scheme," *Fuzzy Sets Syst.*, vol. 138, no. 1, pp. 53–65, 2003.

[7] Y. Chang, S. Chen, and C. Liau, "Fuzzy interpolative reasoning for sparse fuzzy rule-based systems based on the areas of fuzzy sets," *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 5, pp. 1285–1301, Oct. 2008.

[8] S. Chen and Y. Chang, "Fuzzy rule interpolation based on the ratio of fuzziness of interval type-2 fuzzy sets," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 12 202–12 213, 2011.

[9] S. Chen, Y. Chang, and J. Pan, "Fuzzy rules interpolation for sparse fuzzy rule-based systems based on interval type-2 Gaussian fuzzy sets and genetic algorithms," *IEEE Trans. Fuzzy Syst.*, vol. 21, no. 3, pp. 412–425, Jun. 2013.

[10] S. Chen and Y. Ko, "Fuzzy interpolative reasoning for sparse fuzzy rule-based systems based on α-cuts and transformations techniques," *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 6, pp. 1626–1648, Dec. 2008.

[11] R. Diao, F. Chao, T. Peng, N. Snooke, and Q. Shen, "Feature selection inspired classifier ensemble reduction," *IEEE Trans. Cybern.*, 2013, to be published. DOI: 10.1109/TCYB.2013.2281820.

[12] R. Diao and Q. Shen, "Feature selection with harmony search," *IEEE Trans. Syst., Man, Cybern. B*, vol. 42, no. 6, pp. 1509–1523, Dec. 2012.

[13] D. Dubois and H. Prade, "On fuzzy interpolation*," *Int. J. General Syst.*, vol. 28, no. 2-3, pp. 103–114, 1999.

[14] G. Feng, "A survey on analysis and design of model-based fuzzy control systems," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 5, pp. 676–697, Oct. 2006.

[15] X. Fu and Q. Shen, "Fuzzy compositional modeling," *IEEE Trans. Fuzzy Syst.*, vol. 18, no. 4, pp. 823–840, Aug. 2010.

[16] I. Gadaras and L. Mikhailov, "An interpretable fuzzy rule-based classification methodology for medical diagnosis," *Artif. Intell. Med.*, vol. 47, no. 1, pp. 25–41, 2009.

[17] H. A. Hagras, "A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots," *IEEE Trans. Fuzzy Syst.*, vol. 12, no. 4, pp. 524–539, Aug. 2004.

[18] F. Hoffmann, D. Schauten, and S. Holemann, "Incremental evolutionary design of tsk fuzzy controllers," *IEEE Trans. Fuzzy Syst.*, vol. 15, no. 4, pp. 563–577, Aug. 2007.

[19] W. Hsiao, S. Chen, and C. Lee, "A new interpolative reasoning method in sparse rule-based systems," *Fuzzy Sets Syst.*, vol. 93, no. 1, pp. 17–22, 1998.

[20] Z. Huang and Q. Shen, "Fuzzy interpolative reasoning via scale and move transformations," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 2, pp. 340–359, Apr. 2006.

[21] Z. Huang and Q. Shen, "Fuzzy interpolation and extrapolation: A practical approach," *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 1, pp. 13–28, Feb. 2008.

[22] R. Jensen and Q. Shen, "New approaches to fuzzy-rough feature selection," *IEEE Trans. Fuzzy Syst.*, vol. 17, no. 4, pp. 824–838, Aug. 2009.

[23] S. Jin, R. Diao, C. Quek, and Q. Shen, "Backward fuzzy rule interpolation with multiple missing values," in *Proc. Int. Conf. Fuzzy Syst.*, 2013, pp. 1–8.

[24] S. Jin, R. Diao, and Q. Shen, "Backward fuzzy interpolation and extrapolation with multiple multi-antecedent rules," in *Proc. Int. Conf. Fuzzy Syst.*, 2012, pp. 1170–1177.

[25] Y. Jin, "Fuzzy modeling of high-dimensional systems: complexity reduction and interpretability improvement," *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 2, pp. 212–221, Apr. 2000.

[26] T. A. Johansen, R. Shorten, and R. Murray-Smith, "On the interpretation and identification of dynamic Takagi-Sugeno fuzzy models," *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 3, pp. 297–313, Jun. 2000.

[27] L. T. Kóczy, K. Hirota, and L. Muresan, "Interpolation in hierarchical fuzzy rule bases," in *Proc. Int. Conf. Fuzzy Syst.*, 2000, pp. 471–477.

[28] L. Koczy and K. Hirota, "Approximate reasoning by linear rule interpolation and general approximation," *Int. J. Approx. Reason.*, vol. 9, no. 3, pp. 197–225, 1993.

[29] L. Koczy and K. Hirota, "Interpolative reasoning with insufficient evidence in sparse fuzzy rule bases," *Inf. Sci.*, vol. 71, no. 1–2, pp. 169–201, 1993.

[30] L. Koczy and K. Hirota, "Size reduction by interpolation in fuzzy rule bases," *IEEE Trans. Syst., Man, Cybern. B*, vol. 27, no. 1, pp. 14–25, Feb. 1997.

[31] S. Kovács, "Extending the fuzzy rule interpolation "FIVE" by fuzzy observation," *Comput. Intell., Theory Appl.*, vol. 38, pp. 485–497, 2006.

[32] S. Kovács, "Special issue on fuzzy rule interpolation," *J. Adv. Comput. Intell. Intell. Informat.*, p. 253, 2011.

[33] L. Kuncheva, "Fuzzy versus nonfuzzy in combining classifiers designed by boosting," *IEEE Trans. Fuzzy Syst.*, vol. 11, no. 6, pp. 729–741, Dec. 2003.

[34] L. Lee and S. Chen, "Fuzzy interpolative reasoning using interval type-2 fuzzy sets," *New Front. Appl. Artif. Intell.*, vol. 5027, pp. 92–101, 2008.

[35] M. Lee, H. Chung, and F. Yu, "Modeling of hierarchical fuzzy systems," *Fuzzy Sets Syst.*, vol. 138, no. 2, pp. 343–361, 2003.

[36] Z.-Q. Liu and R. Satur, "Contextual fuzzy cognitive map for decision support in geographic information systems," *IEEE Trans. Fuzzy Syst.*, vol. 7, no. 5, pp. 495–507, Oct. 1999.

[37] I. Miguel and Q. Shen, "Fuzzy rrDFCSP and planning," *Artif. Intell.*, vol. 148, no. 1, pp. 11–52, 2003.

[38] N. Naik, R. Diao, C. Quek, and Q. Shen, "Towards dynamic fuzzy rule interpolation," in *Proc. Int. Conf. Fuzzy Syst.*, 2013, pp. 1–7.

[39] N. M. Parthalain and R. Jensen, "Simultaneous feature and instance selection using fuzzy-rough bireducts," in *Proc. Int. Conf. Fuzzy Syst.*, 2013, pp. 1–7.

[40] W. L. Perry, C. Berrebi, R. A. Brown, J. Hollywood, A. Jaycocks, P. Roshan, T. Sullivan, and L. Miyashiro, *Predicting Suicide Attacks: Integrating Spatial, Temporal, and Social Features of Terrorist Attack Targets*. Santa Monica, CA, USA: RAND, 2013.

[41] Q. Shen, J. Keppens, C. Aitken, B. Schafer, and M. Lee, "A scenario-driven decision support system for serious crime investigation," *Law, Probabil. Risk*, vol. 5, no. 2, pp. 87–117, 2006.

[42] A. Tajbakhsh, M. Rahmati, and A. Mirzaei, "Intrusion detection using fuzzy association rules," *Appl. Soft Comput.*, vol. 9, no. 2, pp. 462–469, 2009.

[43] D. Tikk and P. Baranyi, "Comprehensive analysis of a new fuzzy rule interpolation method," *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 3, pp. 281–296, Jun. 2000.

[44] D. Tikk, I. Joó, L. Kóczy, P. Várlaki, B. Moser, and T. Gedeon, "Stability of interpolative fuzzy KH controllers," *Fuzzy Sets Syst.*, vol. 125, no. 1, pp. 105–119, 2002.

[45] A. R. Várkonyi-Kóczy, A. Almos, and T. Kovácsházy, "Genetic algorithms in fuzzy model inversion," in *Proc. Int. Conf. Fuzzy Syst.*, 1999, vol. 3, pp. 1421–1426.

[46] A. R. Várkonyi-Kóczy, G. Péceli, T. P. Dobrowiecki, and T. Kovácsházy, "Iterative fuzzy model inversion," in *Proc. Int. Conf. Fuzzy Syst.*, 1998, vol. 1, pp. 561–566.

[47] D. Waltz, "Understanding line drawings of scenes with shadows," in *The Psychology of Computer Vision*. New York, NY, USA: McGraw-Hill, 1975, pp. 11–91.

[48] D. Wang, X. Zeng, and J. Keane, "Intermediate variable normalization for gradient descent learning for hierarchical fuzzy system," *IEEE Trans. Fuzzy Syst.*, vol. 17, no. 2, pp. 468–476, Apr. 2009.

[49] K. W. Wong, D. Tikk, T. D. Gedeon, and L. T. Kóczy, "Fuzzy rule interpolation for multidimensional input spaces with applications: A case study," *IEEE Trans. Fuzzy Syst.*, vol. 13, no. 6, pp. 809–819, Dec. 2005.

[50] Y. Yam and L. Kóczy, "Representing membership functions as points in high-dimensional spaces for fuzzy interpolation and extrapolation," *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 6, pp. 761–772, Dec. 2000.

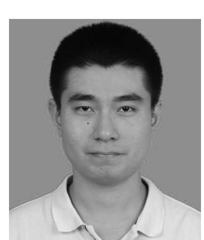
[51] Y. Yam, M. Wong, and P. Baranyi, "Interpolation with function space representation of membership functions," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 3, pp. 398–411, Jun. 2006.

[52] L. Yang and Q. Shen, "Adaptive fuzzy interpolation," *IEEE Trans. Fuzzy Syst.*, vol. 19, no. 6, pp. 1107–1126, Dec. 2011.

[53] L. Yang and Q. Shen, "Closed form fuzzy interpolation," *Fuzzy Sets Syst.*, vol. 225, pp. 1–22, 2013.

Shangzhu Jin received the B.Sc. degree in computer science from Beijing Technology and Business University, Beijing, China, and the M.Sc. degree in control theory and control engineering from Yanshan University, Qinhuangdao, China. He is currently working toward the Ph.D. degree in computer science with the Department of Computer Science, Institute of Mathematics, Physics and Computer Science, Aberystwyth University, Aberystwyth, U.K.

His research interests include fuzzy systems, approximative reasoning, and network security.



Ren Diao received the M.A. and M.Sc. degrees from the University of Cambridge, U.K., and the University of Birmingham, U.K., respectively, and the Ph.D. degree from Aberystwyth University, Aberystwyth, U.K.

He is currently a Research Fellow with the Department of Computer Science, Institute of Mathematics, Physics and Computer Science, Aberystwyth University. His research interests include fuzzy set theory, nature inspired heuristics, and machine learning. He has published more than 20 peer-reviewed papers in leading international journals and conferences.

Chai Quek (SM'10) received the B.Sc. degree in electrical and electronics engineering and the Ph.D. degree in intelligent control from Heriot-Watt University, Edinburgh, U.K., in 1986 and 1990, respectively.

He is currently the Assistant Chair with the School of Computer Engineering, Nanyang Technological University (NTU), Singapore, and a Research Member with the Centre for Computational Intelligence, NTU. His research interests include intelligent control, intelligent architectures, computational finance, neural networks, fuzzy neural systems, neurocognitive informatics, and genetic algorithms.

Dr. Quek is a member of the IEEE Technical Committee on Computational Finance.

Qiang Shen received the Ph.D. degree from Heriot-Watt University, Edinburgh, U.K., and the D.Sc. degree from Aberystwyth University, Aberystwyth, U.K.

He holds the established Chair in computer science and is the Director with the Institute of Mathematics, Physics and Computer Science, Aberystwyth University. His research interests include computational intelligence, reasoning under uncertainty, pattern recognition, data mining, and real-world applications of such techniques for intelligent decision support (e.g., crime detection, consumer profiling, systems monitoring, and medical diagnosis). He has authored two research monographs and more than 310 peer-reviewed papers.

Dr. Shen is a long-serving Associate Editor of the IEEE TRANSACTIONS ON CYBERNETICS and the IEEE TRANSACTIONS ON FUZZY SYSTEMS and an editorial board member of several other leading international journals. He received an Outstanding Transactions Paper Award from IEEE. He is a Fellow of the Learned Society of Wales.