Integration of Interpolation and Inference with Multi-antecedent Rules

Nitin Naik[1] and Qiang Shen[2]

[1] Defence School of Communications and Information Systems,
Ministry of Defence, United Kingdom
[2] Department of Computer Science, Aberystwyth University, United Kingdom
`nitin.naik100@mod.gov.uk, qqs@aber.ac.uk`

Abstract. The efficacious fuzzy rule based systems perform their tasks with either a dense rule base or a sparse rule base. The nature of the rule base decides on whether compositional rule of inference (CRI) or fuzzy rule interpolation (FRI) should be applied. Given a dense rule base where at least one rule exists for every observation, CRI can be effectively and sufficiently employed. For a sparse rule base where rules do not cover all possible observations, FRI is required. Nonetheless, certain observations may be matched partly or completely with any of the existing rules in the sparse rule-base. Such observations can be directly dealt with using CRI and the conclusion can be inferred via firing the matched rule, thereby avoiding extra overheads of interpolation. If no such matching can be found then correct rules should be selected to ensure the accuracy while performing FRI. This paper proposes a generalised approach for the integration of FRI and CRI. It utilises the notion of alpha-cut overlapping to determine the matching degree between rule antecedents and a given observation in order to determine if CRI is to be applied. In the event of no matching rules, the nearest rules will be chosen to derive conclusion using FRI based on the the best suitable distance metric among possible alternatives such as the Centre of Gravity, Hausdorff Distance and Earth Mover's Distance. Comparative results are presented to demonstrate the effectiveness of this integrated approach.

Keywords: Computational Rule of Inference · Rule Interpolation · Rule Extrapolation · Integration of Interpolation and Inference · Multi-antecedent Rules.

1 Introduction

Fuzzy systems infer the results based on the use of either dense or sparse rule bases. For a sparse rule base where rules do not cover possible observations, fuzzy rule interpolation (FRI) is the most popular way of approximating a conclusion [2], [4]. Nonetheless, interpolation generally incurs more overheads for generating approximate conclusions [8], [10]. Despite the sparsity of a rule base, on many occasions, observations may still match with a certain existing rule partly or completely, which may avoid the need of conducting rule interpolation [9], [11]. This can be ensured using a suitable pre-interpolation inference technique such as compositional rule of inference (CRI) [22]. If this pre-interpolation inference technique could not find any matching rule then interpolation would be the next operation to obtain the result [12]. This leads the way for

finding a simple and fast mechanism to determine when to apply CRI or FRI given a sparse rule base [18], [19].

Fuzzy rule interpolation and extrapolation requires only few closest rules to infer results [5]. Therefore, the selection of appropriate closest rules determines the correctness of results generated by interpolation or extrapolation. Consequently, if the selected rules are not realistically the closest ones then the results may be inaccurate irrespective of the FRI approach employed, even though it may be a generally powerful technique. This reveals the fact that the selection of what distance metric to use is critical in finding the correct closest rules for the given observation [6]. Many rule interpolation or extrapolation methods use the most popular distance metric based on the Centre of Gravity (COG) of the membership functions concerned [7]. Unfortunately, the COG values may be the same for two completely different fuzzy sets, which can be seen in Fig. 2. This leads to present further investigation for effective distance metrics in an effort to improve the accuracy of FRI approach, especially those transformation based techniques [3].

The potential efficiency gains of running CRI prior to interpolation and the effective use of distance metrics in FRI have both motivated the development of an integrated solution for inference with fuzzy rule-based systems [13], [15], [16], [17]. This paper presents an integration of interpolation and inference to obtain the best of both for systems with a sparse fuzzy rule base. This integrated approach determines the possibility of applying CRI based on exact or partial matching rule(s) in the sparse rule base, whereby minimising the interpolation overheads [20]. If no matching rule is found in the rule base then certain efficacious distance metric is applied to obtain the best closest rules. Here, alongside with COG, two alternatives, namely the Hausdorff Distance (HD) and the Earth Mover's Distance (EMD) are introduced, which are tested and compared against the use of COG to determine the more correct nearest rules for interpolation/extrapolation. The HD calculates the proximity of two fuzzy sets and provides the scalar score of the similarity between them, while the EMD calculates the similarity between the two multi-dimensional distributions over a region. As a result, HD and EMD distances can find the closest rules for a given observation more precisely than the COG based distance metric. This helps increase the overall accuracy of the integrated approach.

This paper is organised as follows. Section 2 presents an overview of the proposed method for integrating inference and interpolation. It works by exploiting the α-cut threshold in an effort to decide on the suitability of whether to run CRI or FRI, and on the applicability of which preferred distance metric to employ for selecting the closest rules to perform interpolation or extrapolation. Section 3 elucidates the implementation procedure of the integrated system and its components. Section 4 reports on the experimental results with detailed comparison among the use of COG, HD and EMD. Section 5 summarises the proposed approach and discusses relevant further research.

2   Integrating Interpolation and Inference

Rule inference (CRI) and rule interpolation (FRI) have been a vital part of developing fuzzy rule-based systems. The use of any particular method is dependent on the type
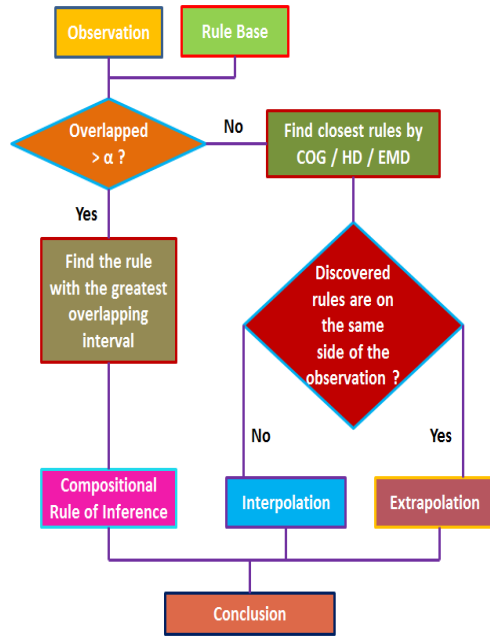
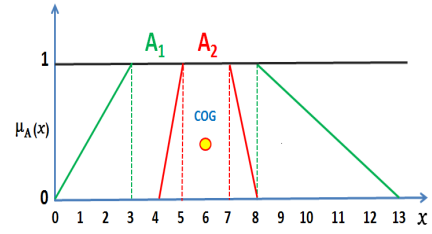Fig. 1: Integrated system with both interpolation and inference

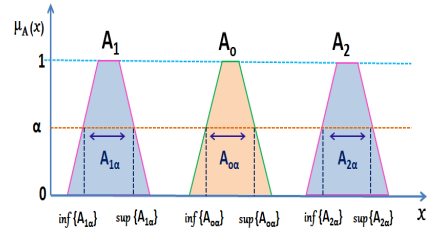Fig. 2: Two dissimilar trapezoidal fuzzy sets of the same COG value

Fig. 3: α-cut concept in trapezoidal fuzzy sets

of the rule base employed in the system. CRI is useful for dense rule base, where an observation matches with any existing rule partly or completely [21]. However, in real-life applications, the design of the dense rule base is expensive or unachievable [14]. In such situations, sparse rule base is the default options and thus FRI. FRI only generates approximate conclusions, which may not be necessarily equally accurate as the CRI result which may be achieved via partial or complete rule matching. The integration of the two methods can provide the best of both whilst compensating the limitations of each other. A particular implementation of this idea is proposed here.

Figure 1 outlines the operational procedure of the proposed integrated system. For any given observation, first it checks the applicability of the CRI by finding a matching rule in the existing rule base based on the use of a pre-determined α-cut threshold. If one or more matching rules are found to be above the pre-determined α-cut threshold, it selects the most overlapped rule amongst all the matched ones and infer the conclusion using CRI. However, if it fails to find any matching rule above the pre-determined α-cut threshold, the interpolation (or extrapolation) becomes the natural choice to infer the conclusion [7]. For improving the results of such interpolation, it utilises one of the possible alternative distance metrics, taken from the set of COG, HD and EMD, to identify the closest rules to the observation. The choice of which distance metric to use depends on its suitability for a given problem domain. From the above, depending on the locations of the closest rules, the system performs the relevant operation of interpolation or extrapolation. In this work, owing to the popularity and availability, the scale

and move transformation interpolation (T-FRI) method [3] is applied to compute the interpolated conclusions. Nevertheless, if desired, any other FRI approach may be used in place of the T-FRI method.

3   Implementation of Proposed Approach

The integrated inference and interpolation system as proposed above has two distinct sub-systems: an α-cut matching component for CRI and a distance metric (COG, HD or EMD) based component for selecting closest rules for interpolation/extrapolation. For this implementation, by following the conventional T-FRI implementations, trapezoidal fuzzy sets are considered. Also, each rule is entitled to involve multiple antecedent variables.

3.1 α-cut Matching for Inference

An α-cut level is used to acquire a crisp set from a fuzzy set based on the pre-determined threshold of α-cut. Suppose that in the universe of discourse X, A is a fuzzy set with the membership function $\mu_A(x)$. Then, for $\alpha \in [0, 1]$, the α-cut of A is [4]:

$$A_\alpha = \{x \in X | \mu_A(x) \geq \alpha\} \tag{1}$$

This notion of α-cuts for two rule antecedents that take on trapezoidal fuzzy sets A_1 and A_2 and that for an observed trapezoidal fuzzy set $A_\circ$ are together shown in Fig. 3, where inf and sup stand for $infimum$ and $supremum$ operators respectively, and α is the given α-cut threshold. In computing α-cut matching, an observation $A_{\circ,j} = (a_{\circ,0}, a_{\circ,1}, a_{\circ,2}, a_{\circ,3})$ over a certain variable X_j is compared with the possible antecedent values $A_{i,j} = (a_{i,j,0}, a_{i,j,1}, a_{i,j,2}, a_{i,j,3})$, based on the given α level, where $i = 1, 2, ..., n$ (indicating that there are n rules which involve x_j) and $j = 1, 2, ..., N$, (indicating that there are N antecedent variables in the problem). Suppose that $[inf\{A_{\circ,j}\}, sup\{A_{\circ,j}\}]$ and $[inf\{A_{i,j}\}, sup\{A_{i,j}\}]$ denote the α-cut of $A_{\circ,j}$ and that of $A_{i,j}$ respectively. Then the check for α-cut matching is simply implemented by assessing whether either of the following holds:

$inf\{A_{\circ,j}\} \leq inf\{A_{i,j}\} \leq sup\{A_{\circ,j}\}$ or $inf\{A_{\circ,j}\} \leq sup\{A_{i,j}\} \leq sup\{A_{\circ,j}\}$.

Finding the matching rules requires computation only above the pre-determined threshold of α-cut. As such, it saves a significant amount of computation for otherwise firing all those rules below the α-cut threshold and improves the efficiency of the system. In particular, if there is only one existing rule matching with the observation with all antecedents above the threshold of α-cut, thereupon, the conclusion is derived from this matched rule using CRI. In case of more than one matched rule, the rule with the greatest degree of matching is selected to derive the conclusion using CRI.

In implementation, the total degree of matching between one rule and an observation is calculated as follows: it sums up all areas constructed by the overlap between the two α-cut fuzzy sets involved in each pair of the fuzzy sets that describe a certain antecedent variable and its counterpart in the observation. The selected rule should have the largest sum of the overlapping areas amongst all the matched rules. Consequently, interpolation is avoided for sparse rule base in such α-cut matching situations. However, if no rule is matched with at least the α-cut level, then fuzzy rule interpolation becomes necessary.

3.2   Closest Rule Selection for Interpolation

When no rule from the existing rule base is matched with the given observation above the α-cut threshold, FRI is the most preferred choice to infer the conclusion. However, its accuracy is dependent on the selected closest rules to derive the conclusion. This highlights the importance of a suitable distance metric to determine the correct closest rules for FRI. Most FRI techniques utilise the centre of gravity (COG) distance metric, which performs well for certain situations but not in many others. When it measures the distances incorrectly, inaccurate or even incorrect interpolation/extrapolation results. To address this issue, in this paper, the Hausdorff Distance (HD) metric and Earth Mover's Distance (EMD) are implemented with FRI as potential alternatives to the COG-based metric. For completeness, more details of these metrics are given below.

Centre of Gravity (COG) COG is computed regarding an imaginary point in a physical object of matter where, for convenience in certain calculations, the total weight of the object is deemed to be concentrated an average of the masses factored by their distances from a unique reference point. The closeness of two fuzzy sets can be determined based on their unique COG points. In particular, given an observation O: $A_{\circ,j} = (A_{\circ,j,1}, A_{\circ,j,2}, A_{\circ,j,3}, A_{\circ,j,4})$ and the antecedent value of the corresponding variable x_j within the ith rule R_i: $A_{i,j} = (A_{i,j,1}, A_{i,j,2}, A_{i,j,3}, A_{i,j,4})$, the COG distance between these two fuzzy sets is defined by:

$$COG(R_i, O) = \sum_{j=1}^{N} \frac{d(COG(A_{\circ,j}), COG(A_{i,j}))}{range_{x_j}} \tag{2}$$

where $COG(A_{\circ,j})$ and $COG(A_{i,j})$ are the COGs of the sets $A_{\circ,j}$ and $A_{i,j}$ respectively, and $range_{x_j} = \max x_j - \min x_j$ is defined over the domain of the variable x_j.

Hausdorff Distance (HD) HD computes the proximity of two arbitrary subsets (sets of points) of a metric space, returning a scalar score of similarity between the two sets of points [1]. It measures the maximum distance of one set $A_{\circ,j}$ from the closest point of the other set $A_{i,j}$. For the proposed work herein, given an observation O: $A_{\circ,j} = (A_{\circ,j,1}, A_{\circ,j,2}, A_{\circ,j,3}, A_{\circ,j,4})$ and the antecedent value of the corresponding variable x_j within the ith rule R_i: $A_{i,j} = (A_{i,j,1}, A_{i,j,2}, A_{i,j,3}, A_{i,j,4})$, the HD metric is defined by:

$$HD(R_i, O) = \sum_{j=1}^{N} \frac{\max\limits_{l \in \{1,2,3,4\}} \left\{ \min\limits_{k \in \{1,2,3,4\}} \{d_{j,kl}(A_{\circ,j,l}, A_{i,j,k})\} \right\}}{range_{x_j}} \tag{3}$$

where $range_{x_j} = \max x_j - \min x_j$, and $d_{j,kl}$ is any conventional distance metric between the two points involved.

Earth Mover's Distance (EMD) EMD calculates the similarity between two multi-dimensional distributions over a region. The EMD is the minimum amount of work required to transform one distribution into another distribution. The distance is measured

by the minimum amount of computation needed to transform a set $A_{i,j}$ to another $A_{\circ,j}$. Here, a unit of computation corresponds to the cost required to calculate the ground distance (i.e., the base distance metric employed as follows) [23]. For this work, given an observation O: $A_{\circ,j} = (A_{\circ,j,1}, A_{\circ,j,2}, A_{\circ,j,3}, A_{\circ,j,4})$ and and the antecedent value of the corresponding variable x_j within the ith rule R_i: $A_{i,j} = (A_{i,j,1}, A_{i,j,2}, A_{i,j,3}, A_{i,j,4})$, the EMD metric is defined by:

$$EMD(R_i, O) = \sum_{j=1}^{N} \frac{\min_{F_j} \frac{\sum_{k=1}^{m} \sum_{l=1}^{n} f_{j,kl} d_{j,kl}}{\sum_{k=1}^{m} \sum_{l=1}^{n} f_{j,kl}}}{range_{x_j}} \tag{4}$$

where $f_{j,kl}$ is the amount of mass transported from $A_{i,j,k}$ to $A_{\circ,j,l}$ for morphing $A_{i,j}$ into $A_{\circ,j}$, and $d_{j,kl}$ is the base distance metric that may be implemented by any standard distance measure.

Depending on the problem domain and computational requirements as discussed next, the best suitable metric out of COG, HD and EMD can be selected and the closest rules discovered for performing interpolation or extrapolation. When all the selected neighbouring rules are on one side of the given observation, extrapolation is carried out to infer the conclusion, if not then interpolation is performed. Here, for simplicity, only two closest rules are considered to perform interpolation or extrapolation using the T-FRI mechanism as per the approach reported in [3]. However, the number of the closest rules needed can be increased if it is desirable for a certain problem domain. The algorithms required to implement the integrated system are shown in Figs. 4-6.

IntegratedSystem(R_i, O, α)
R_i: $A_{i,j} = (A_{i,j,1}, A_{i,j,2}, A_{i,j,3}, A_{i,j,4})$,
O: $A_{\circ,j} = (A_{\circ,j,1}, A_{\circ,j,2}, A_{\circ,j,3}, A_{\circ,j,4})$,
α, α-cut threshold.
  1: $R_{overlap} = \alpha$-CutOverlapping(R_i, O, α);
  2: **if** $R_{overlap} \neq NULL$ **then**
  3:     $B_o = $ CRI($R_{overlap}, O$) ;
  4: **else**
  5:     $R_{close1} = $ COG/HD/EMD_Closest(R, O);
  6:     $R_{close2} = $ COG/HD/EMD_Closest($R - R_{close1}, O$);
  7:     **if** $R_{close1} < O < R_{close2}$ **or** $R_{close2} < O < R_{close1}$ **then**
  8:       $B_o = $ Interpolation($R_{close1}, O, R_{close2}$) ;
  9:     **else**
10:       $B_o = $ Extrapolation($R_{close1}, R_{close2}, O$) ;
11:     **end if**
12: **end if**

Fig. 4: Integrated approach to interpolation and inference

$\alpha-$CutOverlapping(R_i, O, α)
$inf\{A_{\circ,j}\}$, infimum value of crisp set $A_{\circ,j}$
$sup\{A_{\circ,j}\}$, supremum value of crisp set $A_{\circ,j}$

```
 1:  maxArea = 0, maxIndex = −1;
 2:  for each Rᵢ in R do
 3:     if inf{A₀,ⱼ} ≤ inf{Aᵢ,ⱼ} ≤ sup{A₀,ⱼ} or inf{A₀,ⱼ} ≤ sup{Aᵢ,ⱼ} ≤
        sup{A₀,ⱼ} then
 4:        overlap = overlapping area of A₀,ⱼ and Aᵢ,ⱼ above α;
 5:        if overlap > maxArea then
 6:           maxArea = overlap;
 7:           maxIndex = i;
 8:        end if
 9:     end if
10:  end for
11:  if maxIndex == −1 then
12:     return NULL;
13:  else
14:     return R_maxIndex;
15:  end if
```

Fig. 5: α-cut matching approach

3.3   General Node on Choice of Appropriate Distance Metric

The likelihood of successfully selecting the most appropriate distance metric for use in a given application depends on multiple factors. Experimental evaluation is expected in general to reach such a decision. However, COG can be employed if computational complexity is the main consideration for the problem domain. The complexity of COG is $O(nm)$ for the two polygons having n and m vertices respectively, whereas the complexity of HD is $O(nm^2)$. The complexity of using EMD is $O(N^3 log N)$ for an N-bin histogram. Both HD and EMD may be particularly suitable for multidimensional environments to achieve higher accuracy as empirically proven (see later). In addition, EMD is often better for matching perceptual similarity where the ground distance is perceptually meaningful, whilst HD can be more effective and utilised for those applications where the presence of noise or occlusion is significant.

4   Illustrative and Experimental Case Studies

This section shows the experimental results for the integrated approach. This is based on the use of a sparse rule base, as shown in Table 1 consisting of eight rules each involving four antecedents. The confidence (i.e. α-cut) level $\alpha = 0.5$ is utilised for α-cut matching to check the suitability of CRI and also, for comparing the performances of COG, HD and EMD in their respective action for identifying the best neighbouring rules to perform either interpolation or extrapolation.

COG/HD/EMD_Closest(R, O)

```
1: closeDist = Max_Value, closeIndex = −1;
2: for each Rᵢ in R do
3:    dist = COG/HD/EMD(Rᵢ, O);
4:    if dist < closeDist then
5:       closeDist = dist;
6:       closenIndex = i;
7:    end if
8: end for
9: return R_closeIndex;
```

<div align="center">Fig. 6: COG/HD/EMD distance metric for finding the closet rules</div>

4.1 α-Cut Overlapping Operation for Inference

This case demonstrates how the α-cut overlapping procedure is used to determine the applicability of CRI for a sparse rule base, avoiding the use of FRI and saving computational overheads. Table 4 shows the considered five observations, which overlap with the existing rules of the sparse rule base. The first observation O_1: $A_{\circ,1} = (12.6, 14.3, 15.6, 16.7)$, $A_{\circ,2} = (14.6, 16.3, 17.6, 18.7)$, $A_{\circ,3} = (16.6, 18.3, 19.6, 20.7)$, and $A_{\circ,4} = (19.6, 11.3, 22.6, 23.7)$, overlaps with the existing rule-5 to rule-8 above the α-cut threshold. Where, rule-5 has the least and rule-7 has the most overlapping area with this observation. Therefore, rule-7 is chosen to infer the conclusion using CRI.

Similarly, the second observation O_2: $A_{\circ,1} = (3.8, 4.9, 5.9, 7.0)$, $A_{\circ,2} = (5.8, 6.9, 7.9, 9.0)$, $A_{\circ,3} = (8.8, 9.9, 10.9, 12.0)$, and $A_{\circ,4} = (10.8, 11.9, 12.9, 14.0)$, overlaps with the existing rule-1 to rule-4 above the α-cut threshold. Where, rule-1 has the least and rule-4 has the most overlapping area with this observation. Therefore, rule-4 is chosen for firing to infer the conclusion using CRI. The third observation O_3: $A_{\circ,1} = (11.2, 12.3, 13.2, 13.7)$, $A_{\circ,2} = (13.2, 14.3, 15.2, 15.7)$, $A_{\circ,3} = (16.2, 17.3, 18.2, 18.7)$, and $A_{\circ,4} = (18.2, 19.3, 20.2, 20.7)$, overlaps with the existing rule-5 to rule-8 above the α-cut threshold. Where, rule-7 has the least and rule-5 has the most overlapping area with this observation. Therefore, rule-5 is chosen to infer the conclusion using CRI. For the fourth observation O_4: $A_{\circ,1} = (2.5, 3.8, 4.7, 7.3)$, $A_{\circ,2} = (4.5, 5.8, 6.7, 9.3)$, $A_{\circ,3} = (7.5, 8.8, 9.7, 12.3)$, and $A_{\circ,4} = (9.5, 10.8, 11.7, 14.3)$, overlaps with the existing rule-1 to rule-4 above the α-cut threshold. Where, rule-3 has the least and rule-1 has the most overlapping area with this observation. Therefore, rule-1 is chosen for firing to derive the conclusion using CRI. The final and fifth observation O_5: $A_{\circ,1} = (11.6, 13.1, 14.5, 15.4)$, $A_{\circ,2} = (13.6, 15.1, 16.5, 17.4)$, $A_{\circ,3} = (16.6, 18.1, 19.5, 20.4)$, and $A_{\circ,4} = (17.1, 18.6, 20.0, 20.9)$, overlaps with the existing rule-5 to rule-8 above the α-cut threshold. Where, rule-5 has the least and rule-8 has the most overlapping area with this observation. Therefore, rule-8 is chosen to infer the conclusion using CRI.

4.2 Closest Rules Selection Using COG, HD or EMD

This experiment exhibits and compares the performances of using COG, HD or EMD for closest rules selection. Table 2 shows all given observations and their selected clos-

Table 1: Sparse Fuzzy Rule Base

No.	Antecedents $R\{A_{i,1}, A_{i,2}, A_{i,3}, A_{i,4}\}$	Consequent B_i
$R1$	(2.4, 3.7, 4.6, 7.1), (4.4, 5.7, 6.6, 9.1), (7.4, 8.7, 9.6, 12.1), (9.4, 10.7, 11.6, 14.1)	(5.9, 7.2, 8.1, 10.6)
$R2$	(3.1, 3.7, 4.8, 7.0), (5.1, 5.7, 6.8, 9.0), (8.1, 8.7, 9.8, 12.0), (10.1, 10.7, 11.8, 14.0)	(6.6, 7.2, 8.3, 10.5)
$R3$	(4.3, 4.6, 5.8, 6.8), (6.3, 6.6, 7.8, 8.8), (9.3, 9.6, 10.8, 11.8), (11.3, 11.6, 12.8, 13.8)	(7.8, 8.1, 9.3, 10.3)
$R4$	(3.5, 4.8, 6.1, 6.9), (5.6, 6.8, 8.1, 8.9), (8.6, 9.8, 11.1, 11.9), (10.6, 11.8, 13.1, 13.9)	(7.1, 8.3, 9.6, 10.4)
$R5$	(11.9, 12.5, 13.4, 14.0), (13.9, 14.5, 15.4, 16.0), (16.9, 17.5, 18.4, 19.0), (18.9, 19.5, 20.4, 21.0)	(15.4, 16.0, 16.9, 17.5)
$R6$	(11.8, 13.2, 14.1, 14.8), (13.8, 15.2, 16.1, 16.8), (16.8, 18.2, 19.1, 19.8), (18.8, 20.2, 21.1, 21.8)	(15.3, 16.7, 17.6, 18.3)
$R7$	(11.5, 14.4, 15.2, 16.0), (13.5, 16.4, 17.2, 18.0), (16.5, 19.4, 20.2, 21.0), (18.5, 21.4, 22.2, 23.0)	(15.3, 17.7, 18.9, 19.5)
$R8$	(11.7, 13.1, 14.4, 15.3), (13.7, 15.1, 16.4, 17.3), (16.7, 18.1, 19.4, 20.3), (18.7, 20.1, 21.4, 22.3)	(15.2, 16.6, 17.9, 18.8)

Table 2: Two Closest Rules Determined by COG, HD and EMD

No.	Observation $O\{A_{\circ,1}, A_{\circ,2}, A_{\circ,3}, A_{\circ,4}\}$	Closest Rules by COG	Closest Rules by HD	Closest Rules by EMD
O_1	(7.1, 8.4, 9.8, 11.6), (9.1, 10.4, 11.8, 13.6), (12.1, 13.4, 14.8, 17.6), (14.1, 15.4, 16.8, 19.6)	$R5$ $R3$	$R7$ $R1$	$R7$ $R4$
O_2	(7.1, 7.8, 8.6, 9.5), (9.1, 9.8, 10.6, 11.5), (12.1, 12.8, 13.6, 14.5), (14.1, 14.8, 15.6, 16.5)	$R3$ $R4$	$R1$ $R2$	$R4$ $R2$
O_3	(7.7, 8.6, 9.6, 11.3), (9.7, 10.6, 11.6, 13.3), (12.7, 13.6, 14.6, 16.3), (14.7, 15.6, 16.6, 18.3)	$R5$ $R3$	$R7$ $R8$	$R7$ $R4$
O_4	(9.7, 10.6, 11.3, 12.0), (11.7, 12.6, 13.3, 14.0), (14.7, 15.6, 16.3, 17.0), (16.7, 17.6, 18.3, 19.0)	$R6$ $R8$	$R7$ $R8$	$R5$ $R6$
O_5	(8.4, 8.9, 9.8, 10.3), (10.4, 10.9, 11.8, 12.3), (13.4, 13.9, 14.8, 15.3), (15.4, 15.9, 16.8, 17.3)	$R5$ $R3$	$R7$ $R1$	$R2$ $R7$
O_6	(8.0, 8.5, 9.0, 9.7), (10.0, 10.5, 11.0, 11.7), (13.0, 13.5, 14.0, 14.7), (15.0, 15.5, 16.0, 16.7)	$R3$ $R4$	$R1$ $R2$	$R4$ $R7$
O_7	(7.5, 8.5, 9.0, 10.3), (9.5, 10.5, 11.0, 12.3), (12.5, 13.5, 14.0, 15.3), (14.5, 15.5, 16.0, 17.3)	$R3$ $R4$	$R1$ $R2$	$R4$ $R3$
O_8	(8.1, 8.8, 9.6, 10.3), (10.1, 10.8, 11.6, 12.3), (13.1, 13.8, 14.6, 15.3), (15.1, 15.8, 16.6, 17.3)	$R5$ $R3$	$R1$ $R2$	$R2$ $R1$
O_9	(8.8, 9.3, 10.1, 10.5), (10.8, 11.3, 12.1, 12.5), (13.8, 14.3, 15.1, 15.5), (15.8, 16.3, 17.1, 17.5)	$R5$ $R6$	$R7$ $R8$	$R7$ $R2$
O_{10}	(8.6, 9.0, 9.5, 10.6), (10.6, 11.0, 11.5, 12.6), (13.6, 14.0, 14.5, 15.6), (15.6, 16.0, 16.5, 17.6)	$R5$ $R3$	$R7$ $R8$	$R7$ $R6$

est rules for interpolation or extrapolation, based on these three distance metrics. Here, the results for most of the observations are completely different for the three distance metrics. Surprisingly, for the first observation O_1: $A_{\circ,1} = (7.1, 8.4, 9.8, 11.6)$, $A_{\circ,2} = (9.1, 10.4, 11.8, 13.6)$, $A_{\circ,3} = (12.1, 13.4, 14.8, 17.6)$, and $A_{\circ,4} = (14.1, 15.4, 16.8, 19.6)$, the closest rule (i.e., rule-7) selected by HD and EMD is the furthest rule found by COG as shown in Table 3. The HD and EMD metrics have determined most of the closest rules, where the shapes of the membership functions of their antecedent variables are quite similar to their counterparts in the given observation. This is helpful in maintaining the interpretability of the integrated system.

In this experiment, following conventional T-FRI approaches and empirical results obtained elsewhere [5], only two nearest rules are chosen by either of COG, HD and EMD distance metrics to perform interpolation/extrapolation. However, more than two nearest rules can also be chosen if preferred. For all the observations, these three distance metrics could not find one similar result. For every observation, these distance metrics selected different nearest rules, which lead to different inference results. Interestingly, the experiment shows a distinctive pattern of selecting the nearest rules by every distance metric. The use of COG has regularly chosen rule-3, rule-4 and rule-5; that of HD has regularly chosen rule-1, rule-2 and rule-7; and that of EMD has regularly chosen rule-7, rule-4 and rule-2. Also, examining the outcomes of using EMD, it can

Table 3: Distance Measures between Rule and Observation Using COG, HD and EMD

Observation	Rules	Order of Rules by COG	Order of Rules by HD	Order of Rules by EMD
O	R	$d_{COG}(R_i, O)$	$d_{HD}(R_i, O)$	$d_{EMD}(R_i, O)$
	R1	4.78	4.50	3.51
	R2	4.57	4.60	3.34
$(7.1, 8.4, 9.8, 11.6)$,	R3	3.85	4.80	3.58
$(9.1, 10.4, 11.8, 13.6)$,	R4	3.88	4.69	3.21
$(12.1, 13.4, 14.8, 17.6)$,	R5	**3.72**	4.80	3.40
$(14.1, 15.4, 16.8, 19.6)$	R6	4.25	4.70	3.33
	R7	5.05	**4.40**	**2.89**
	R8	4.40	4.60	3.82

Table 4: Results for α-Cut Matching for Inference

Observation $O\{A_{o,1}, A_{o,2}, A_{o,3}, A_{o,4}\}$	Best Rule
$(12.6, 14.3, 15.6, 16.7)$, $(14.6, 16.3, 17.6, 18.7)$, $(16.6, 18.3, 19.6, 20.7)$, $(19.6, 11.3, 22.6, 23.7)$	$R7$
$(3.8, 4.9, 5.9, 7.0)$, $(5.8, 6.9, 7.9, 9.0)$, $(8.8, 9.9, 10.9, 12.0)$, $(10.8, 11.9, 12.9, 14.0)$	$R4$
$(11.2, 12.3, 13.2, 13.7)$, $(13.2, 14.3, 15.2, 15.7)$, $(16.2, 17.3, 18.2, 18.7)$, $(18.2, 19.3, 20.2, 20.7)$	$R5$
$(2.5, 3.8, 4.7, 7.3)$, $(4.5, 5.8, 6.7, 9.3)$, $(7.5, 8.8, 9.7, 12.3)$, $(9.5, 10.8, 11.7, 14.3)$	$R1$
$(11.6, 13.1, 14.5, 15.4)$, $(13.6, 15.1, 16.5, 17.4)$, $(16.6, 18.1, 19.5, 20.4)$, $(17.1, 18.6, 20.0, 20.9)$	$R8$

be seen that employing EMD yields a greater sensitivity than utilising the other two, as it changes the nearest rules with any minor change in the value of the observation.

4.3   Selection of Interpolation/Extrapolation Mechanism

There are a range of variations in selecting the type of inference method in response to the use of different distance metrics. Indeed, while one selects interpolation and the others may select extrapolation as illustrated in Table 6. Interestingly, however, despite different nearest rules may be chosen, the selected inference operation can be the same for almost half of the given observations. Specially, the use of COG or EMD appears somewhat similar in signifying the type of inference method. Nonetheless, HD is rather different from COG and EMD as its use leads to two interpolation and eight extrapolation operations.

4.4   Accuracy of Interpolation/Extrapolation Results

Table 5 shows the interpolation/extrapolation results with respect to all the given observations using T-FRI based on COG, HD and EMD. These results are further compared against the underlying ground truth values, as reflected in Table 7. Thus, those closest rules selected by HD have produced the best results (over this set of observations). These outcomes confirm the intuition that the use of a carefully chosen distance metric can help improve (or otherwise, affect adversely) the interpolation/extrapolation result. However, it requires further evaluation of the HD and EMD distance metrics on different types of rule and other observations, in order to better verify the accuracy and consistency of their use.

5   Conclusions

This paper has presented an integrated system to perform interpolation and inference effectively and efficiently. Initially, it performs a pre-interpolation inference assessment

Table 5: Interpolation/Extrapolation Results

Obs. O_i	Result Based on COG $Consequent_{COG}(B_{o,i})$	Result Based on HD $Consequent_{HD}(B_{o,i})$	Result Based on EMD $Consequent_{EMD}(B_{o,i})$
O_1	(10.66, 12.81, 12.16, 14.16)	(10.73, 11.97, 13.31, 15.09)	(10.73, 11.99, 13.29, 15.07)
O_2	(10.62, 11.78, 11.54, 13.02)	(10.60, 11.3, 12.1, 13.0)	(10.62, 11.67, 11.66, 13.02)
O_3	(11.27, 12.75, 12.15, 14.77)	(11.36, 12.57, 12.53, 14.74)	(11.36, 12.58, 12.52, 14.74)
O_4	(13.18, 14.51, 14.45, 15.48)	(13.29, 14.17, 14.81, 15.52)	(13.18, 14.55, 14.42, 14.48)
O_5	(11.90, 12.40, 13.30, 13.8)	(11.99, 12.46, 13.33, 13.83)	(11.99, 13.46, 13.33, 13.83)
O_6	(11.52, 12.33, 12.09, 13.22)	(11.5, 12.0, 12.5, 13.2)	(11.59, 12.28, 12.24, 13.23)
O_7	(11.03, 12.55, 11.84, 13.83)	(11.01, 11.58, 12.14, 13.81)	(11.03, 12.55, 11.84, 13.83)
O_8	(11.51, 12.53, 12.51, 13.80)	(11.60, 12.30, 13.10, 13.80)	(10.1, 10.8, 11.6, 12.3)
O_9	(12.30, 12.80, 13.60, 14.0)	(12.38, 12.86, 13.63, 14.02)	(12.40, 12.86, 13.63, 14.03)
O_{10}	(12.07, 12.81, 12.40, 14.17)	(12.23, 12.75, 12.68, 14.16)	(12.23, 12.74, 12.71, 14.15)

Table 6: Suggested Inference Operation Using COG, HD or EMD

Obs. O_i	Inference Method Based on COG	Inference Method Based on HD	Inference Method Based on EMD
O_1	$Interpolation$	$Interpolation$	$Interpolation$
O_2	$Extrapolation$	$Extrapolation$	$Extrapolation$
O_3	$Interpolation$	$Extrapolation$	$Interpolation$
O_4	$Extrapolation$	$Extrapolation$	$Extrapolation$
O_5	$Interpolation$	$Interpolation$	$Interpolation$
O_6	$Extrapolation$	$Extrapolation$	$Interpolation$
O_7	$Extrapolation$	$Extrapolation$	$Extrapolation$
O_8	$Interpolation$	$Extrapolation$	$Extrapolation$
O_9	$Extrapolation$	$Extrapolation$	$Interpolation$
O_{10}	$Interpolation$	$Extrapolation$	$Extrapolation$

Table 7: Accuracy of Interpolation/Extrapolation in Relation to the Use of COG, HD and EMD

Metrics	Ground Truth vs. Result Based on COG $\epsilon_{\%COG}$	Ground Truth vs. Result Based on HD $\epsilon_{\%HD}$	Ground Truth vs. Result Based on EMD $\epsilon_{\%EMD}$
AVG	5.38	1.87	3.62
SD	3.46	2.26	2.34

to determine whether to use compositional rule of inference (CRI). For this, it utilises the α-cut operation to determine the applicability of CRI despite the sparsity of the given rule base. This helps avoid any unnecessary interpolation/extrapolation while attaining better accuracy. Consequently, interpolation/extrapolation is only applied when no match is found for the given observation. For enhancing the accuracy of interpolated/extrapolated results, this paper has introduced a method which utilises the HD or EMD metric to decide on the nearest rules for interpolation or extrapolation. Both HD and EMD metrics are very effective in multidimensional environment, facilitating efficient computation of the required distance measures between values represented by fuzzy sets. Experimental results have shown that the use of either HD or EMD leads to moderately better results for this implementation than the utilisation of the conventional COG-based metric. In future, it is worthwhile to investigate the effectiveness of this approach for a large rule base and multiple rule selection for interpolation/extrapolation.

References

1. Chaudhuri, B.B., Rosenfeld, A.: A modified Hausdorff Distance between fuzzy sets. Information Sciences **118**, 159–171 (1999)
2. Dubois, D., Prade, H.: On fuzzy interpolation. In: 3rd International Conference on Fuzzy Logic & Neural Networks. pp. 353–354 (1995)

3. Huang, Z., Shen, Q.: Fuzzy interpolative reasoning via scale and move transformations. IEEE Transactions on Fuzzy Systems **14**(2), 340–359 (2006)
4. Koczy, L.T., Hirota, K.: Rule interpolation by α-level sets in fuzzy approximate reasoning. J. BUSEFAL, Automne, URA-CNRS **46**, 115–123 (1991)
5. Li, F., Li, Y., Shang, C., Shen, Q.: Fuzzy knowledge-based prediction through weighted rule interpolation. IEEE Transactions on Cybernetics (2019)
6. McCulloch, J.C., Hinde, C.J., Wagner, C., Aickelin, U.: A fuzzy directional distance measure. In: 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) (2014)
7. Naik, N.: Dynamic Fuzzy Rule Interpolation. Ph.D. thesis, Department of Computer Science, Institute of Mathematics, Physics and Computer Science, Aberystwyth University, UK (2015)
8. Naik, N., Diao, R., Quek, C., Shen, Q.: Towards dynamic fuzzy rule interpolation. In: IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). pp. 1–7 (2013)
9. Naik, N., Diao, R., Shang, C., Shen, Q., Jenkins, P.: D-FRI-WinFirewall: Dynamic fuzzy rule interpolation for windows firewall. In: 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) (2017)
10. Naik, N., Diao, R., Shen, Q.: Genetic algorithm-aided dynamic fuzzy rule interpolation. In: IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). pp. 2198–2205 (2014)
11. Naik, N., Diao, R., Shen, Q.: Application of dynamic fuzzy rule interpolation for intrusion detection: D-FRI-Snort. In: IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). pp. 78–85 (2016)
12. Naik, N., Diao, R., Shen, Q.: Dynamic fuzzy rule interpolation and its application to intrusion detection. IEEE Transactions on Fuzzy Systems **26**(4), 1878–1892 (2018)
13. Naik, N., Jenkins, P.: Enhancing windows firewall security using fuzzy reasoning. In: IEEE International Conference on Dependable, Autonomic and Secure Computing (2016)
14. Naik, N., Jenkins, P.: Fuzzy reasoning based windows firewall for preventing denial of service attack. In: IEEE International Conference on Fuzzy Systems. pp. 759–766 (2016)
15. Naik, N., Jenkins, P.: A fuzzy approach for detecting and defending against spoofing attacks on low interaction honeypots. In: 21st International Conference on Information Fusion. pp. 904–910. IEEE (2018)
16. Naik, N., Jenkins, P., Cooke, R., Ball, D., Foster, A., Jin, Y.: Augmented windows fuzzy firewall for preventing denial of service attack. In: 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) (2017)
17. Naik, N., Jenkins, P., Kerby, B., Sloane, J., Yang, L.: Fuzzy logic aided intelligent threat detection in cisco adaptive security appliance 5500 series firewalls. In: 2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) (2018)
18. Naik, N., Shang, C., Shen, Q., Jenkins, P.: Intelligent dynamic honeypot enabled by dynamic fuzzy rule interpolation. In: The 4th IEEE International Conference on Data Science and Systems (DSS-2018). pp. 1520–1527. IEEE (2018)
19. Naik, N., Shang, C., Shen, Q., Jenkins, P.: Vigilant dynamic honeypot assisted by dynamic fuzzy rule interpolation. In: IEEE Symposium Series on Computational Intelligence (2018)
20. Naik, N., Shang, C., Shen, Q., Jenkins, P.: D-FRI-CiscoFirewall: Dynamic fuzzy rule interpolation for Cisco ASA Firewall. In: IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). IEEE (2019)
21. Naik, N.: Fuzzy inference based intrusion detection system: FI-Snort. In: IEEE International Conference on Dependable, Autonomic and Secure Computing. pp. 2062–2067 (2015)
22. Naik, N., Su, P., Shen, Q.: Integration of interpolation and inference. In: UK Workshop on Computational Intelligence. pp. 1–7 (2012)
23. Rubner, Y., Tomasi, C., Guibas, L.J.: A metric for distributions with applications to image databases. In: IEEE International Conference on Computer Vision. pp. 59–66 (1998)