

Tuning Larger Membership Grades for Fuzzy Association Rules

Stephen G. Matthews

Abstract—Sigma count measures scalar cardinality of fuzzy sets. A problem with sigma count is that values of scalar cardinality are calculated entirely from many small membership grades or entirely from few large membership grades. Two novel scalar cardinality measures are proposed for the fitness of a genetic algorithm for tuning membership functions prior to fuzzy association rule mining so that individual membership grades are larger. Preliminary results show a decrease in small membership grades and an increase in large membership grades for fuzzy association rules tested on real-world benchmark datasets.

I. INTRODUCTION

Association rule mining is an exploratory and descriptive rule induction process of identifying significant relations between items in Boolean transaction datasets [1] used for data analysis and interpretation. Association rules are expressed as an implication of the form $X \Rightarrow Y$ where the antecedent and consequent are sets of Boolean items, which are known as itemsets, where $X \cap Y = \emptyset$. For example, $\text{beer} \Rightarrow \text{pizza}$ means there is a frequent association between customers who purchased beer and pizza.

For datasets containing numeric items instead of Boolean items, quantitative association rules [2] model the relations between intervals of items containing numeric values, such as 4–5 items purchased, greater than 60 seconds spent visiting a Web page, or company shares greater than 1,000 [2]. In fuzzy association rule mining, concepts are described with linguistic terms, and fuzzy sets [3] model the uncertainty and imprecision of linguistic terms used by humans.

The support-confidence framework introduced two measures for Boolean association rules [4]. Support measures the number of transactions for a rule $X \Rightarrow Y$ in which both X and Y occur. Confidence measures the conditional probability of Y being in a transaction given that it contains X . The generalisation of support and confidence from crisp sets to fuzzy sets often uses Zadeh's sigma count σ for fuzzy support

$$\text{FuzzySup}(A \Rightarrow B) = \sigma(\min(A, B)), \quad (1)$$

where the minimum \min is used for the intersection of two fuzzy sets A and B [3].

Fuzzy support of an itemset can also be expressed as scalar cardinality of a fuzzy set containing dataset transactions [5]. The membership grades express the degree of belonging of transactions to an itemset. For example, consider a dataset of 19 transactions and 2 itemsets C and D

Stephen G. Matthews is with the Intelligent Systems Laboratory, Department of Engineering Mathematics, University of Bristol, Bristol BS8 1UB, UK (email: stephen.matthews@bristol.ac.uk).

$$C = \left\{ \frac{x_1}{0.9}, \frac{x_2}{1.0} \right\},$$
$$D = \left\{ \frac{x_1}{0.1}, \frac{x_2}{0.1}, \frac{x_3}{0.1}, \dots, \frac{x_{19}}{0.1} \right\}.$$

The sigma count causes a problem in the scalar cardinality of a fuzzy set [6] [7]. The scalar cardinality of each set using the sigma count is 1.9. However, A contains only 2 members that both have large membership grades, and B contains 19 members with small membership grades. The fuzzy supports of A and B are the same when the number of transactions is very different. The sigma count does not distinguish between many elements with small membership grades and few elements with large membership grades.

The sigma count problem effects the accuracy and interpretability of fuzzy association rules, and more generally fuzzy rule-based systems. Consider two fuzzy association rules from C and D in the previous example. Many small membership grades suggests the rule is a less accurate description of a relation in the dataset, and also the meaning and interpretability of the rule is inconsistent.

This paper proposes two scalar cardinality measures of fuzzy sets of transactions, which are based on the sigma count, for tuning the membership functions (of fuzzy sets describing linguistic terms of variables) prior to mining fuzzy association rules. The aim of the proposed scalar cardinality measures is to guide the genetic tuning towards more larger membership grades in the fuzzy set.

The outline of the paper is as follows: Section II reviews previous related work, Section III discusses the effect of the sigma count problem on the accuracy-interpretability problem, Section IV presents the method of tuning the membership functions and proposes the two scalar cardinality measures of fuzzy sets, Section V presents preliminary results demonstrating the capability of the proposal, and Section VI draws conclusions.

II. RELATED WORK

Previous research on fuzzy association rule mining, the sigma count problem, and cardinality of fuzzy sets are reviewed.

Fuzzy sets were initially proposed as a less terse summarisation of data with linguistic terms in data mining [8]. After the concept of fuzzy association rule induction was initially proposed without an algorithm [9], the F-APACS algorithm for fuzzy association rule mining was proposed [10]. FuzzyApriori [11] is an extension of Apriori [4] for fuzzy association rule mining. FuzzyApriori performs the same bottom-up search as Apriori combined with the downward closure property for discovering itemsets (a subset of

all items in the dataset) greater than or equal to the minimum fuzzy support and rules are generated that are greater than or equal to the minimum confidence. Fuzzy association rules have been used for descriptive data mining, such as Web usage mining [12], and also predictive data mining [13].

As described in the introduction, the sigma count problem affects support but it also affects confidence. The following example [14] demonstrates this. Confidence is measured as

$$\text{Conf}(A \Rightarrow B) = \frac{\sigma(\min(A, B))}{\sigma(A)}, \quad (2)$$

for two fuzzy sets A and B containing different members

$$A = \left\{ \frac{x_1}{1} \right\},$$

$$B = \left\{ \frac{x_2}{1} \right\},$$

which leads to $\text{Conf}(A, B) = 0$. However,

$$A = \left\{ \frac{x_1}{1.0}, \frac{x_2}{0.01}, \frac{x_3}{0.01}, \dots, \frac{x_{1000}}{0.01} \right\},$$

$$B = \left\{ \frac{x_1}{0.01}, \frac{x_2}{1.0}, \frac{x_3}{0.01}, \dots, \frac{x_{1000}}{0.01} \right\},$$

leads to

$$\text{Conf}(A, B) = \frac{1000 \times 0.01}{1 + 999 \times 0.01} \approx 0.91,$$

which is large for two disjoint sets.

The cardinality of a fuzzy set can be measured with either scalar cardinality by aggregating integer or real values, or fuzzy cardinality where the measure is a fuzzy set. The sigma count [3] and power of a fuzzy set [15] are scalar cardinalities that sum membership grades. To alleviate the effects of the sigma count problem, the cardinality of the α -cut at levels of 0.5 has been used to replace real values with integer values [16]. A similar approach proposed an interval between the cardinality of the α -cut at 0.5 and the cardinality of the *strong* α -cut at level 0.5 [17]. A different integer interval was proposed in the range of the lower and upper expectations of cardinality [18]. Weighted cardinality was proposed [14], and a weighted summation of the cardinalities of alpha cuts was also proposed [19].

A new, alternative perspective is to view fuzzy cardinality for association rules as fuzzy numbers [6]. Different interpretations of fuzzy sets have recently been proposed that offer alternatives for dealing with the impact of small membership values. Instead of mapping a universal set of elements onto a membership scale, the $X-\mu$ approach maps from the membership scale to the universe, and allows a visual interpretation of fuzzy cardinality and fuzzy arithmetic [20]. The $X-\mu$ approach has been applied in a fuzzy association rule method [21]. A similar idea is used in gradual numbers and the Representation-by-level (RL) approach, which represents fuzzy concepts by an assignment of crisp sets to levels of fulfilment of a property [22].

III. ACCURACY-INTERPRETABILITY

The accuracy-interpretability problem is a trade-off in the design of a fuzzy rule-based system (FRBS) between creating a true model of a system that is accurate [23] and creating a model that is understandable and interpretable by humans [24] [25]. The problem is also relevant to fuzzy association rule mining because it shares components with FRBSs.

Accuracy is easily measured in FRBSs with measures such as mean square error, so attention has focused on interpretability [26]. For measuring the accuracy of fuzzy association rules, previous research has focused specifically on overcoming the sigma count problem, which has been reviewed in Section II. However, this has not been addressed from the perspective of the accuracy-interpretability problem.

How accurately does a rule represent a relation in a dataset when the rules has many small membership grades compared to another rule with one large membership grade? A rule with many small membership grades is a less accurate description because the scalar cardinality measure does not model each individual item in the count.

The sigma count problem occurs in support measures or scalar cardinality, both of which measure accuracy. This problem also raises questions about the interpretability of fuzzy association rules. Consider two rules with the same fuzzy support, it is not known whether both rules have many small membership grades or few large membership grades. Many small membership grades can be misleading and counter-intuitive because the membership grades of each element are small but the accumulated membership grades are relatively large. Are many small membership grades useful in interpreting fuzzy association rules?

Accuracy and interpretability are competing objectives when learning/tuning FRBSs and membership functions for fuzzy association rules, so a trade-off is often sort. Takagi-Sugeno-Kang FRBSs [27] [28] use fuzzy numbers for approximation, which do not have linguistic terms assigned to fuzzy sets, to achieve better levels of accuracy but at the cost of losing interpretability. Improving accuracy of scalar cardinality is an approach that can maintain the interpretability of linguistic terms assigned to fuzzy sets in descriptive Mamdani approaches.

IV. PROPOSED METHOD

The aim is to improve the accuracy of fuzzy association rules by increasing the number of larger membership grades, whilst maintaining the interpretability of membership functions. A Genetic Algorithm (GA) [29] performs genetic tuning [30] of the membership functions before fuzzy association rules are mined. Our proposed approach considers the data as *fuzzy* transactions in the GA with scalar cardinality and as *non-fuzzy* transactions when mining fuzzy association rules with fuzzy support according to FuzzyApriori.

The proposed method uses an existing GA model for tuning membership functions and the general model of fuzzy association rules and fuzzy transactions, which are both discussed here. The two scalar cardinality measures of fuzzy transactions are proposed for the fitness function of the GA.

A. Genetic Algorithm and 2-tuple linguistic representation

An existing GA was chosen to demonstrate the two scalar cardinality measures. Alcalá-Fdez et al. [31] used the Cross-generational elitist selection, Heterogeneous re-combination, and Cataclysmic mutation (CHC) GA [32] and the 2-tuple linguistic representation [33]. The combination of this GA and fuzzy representation has a proven ability in achieving both accuracy and interpretability of fuzzy association rules [31] [12].

The 2-tuple linguistic representation is based on a symbolic translation of a fuzzy set [33]. The symbolic translation is the lateral displacement of a fuzzy set within the interval $[-0.5, 0.5)$ about the fuzzy set's index and between two linguistic terms on the universe of discourse. The membership function maintains its shape whilst it is laterally displaced left or right from its original membership function but not further than the middle points of adjacent fuzzy sets. The 2-tuple linguistic representation is defined as

$$\{(s_j, \alpha_j) \mid s_j \in S, \alpha_j \in [-0.5, 0.5)\}, \quad (3)$$

where S is a set of linguistic labels, α quantifies the lateral displacement of a linguistic label s_j within the interval $[-0.5, 0.5)$ and j is the index. An example of the 2-tuple linguistic representation is shown in Figure 1.

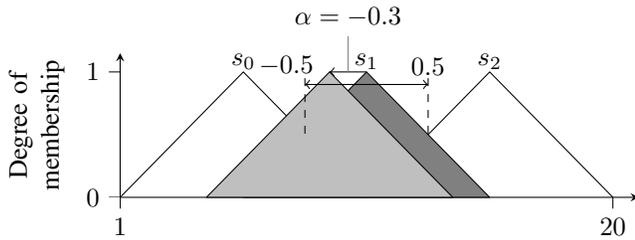


Fig. 1: Example of a 2-tuple membership function $(s_1, -0.3)$ (light grey) that is displaced from membership function s_1 (dark grey)

The CHC algorithm tunes the membership functions by learning the lateral displacements of the 2-tuple linguistic representation. The benefit of this approach is that the accuracy is tuned whilst maintaining interpretability and the dimensionality of the search space is reduced (membership functions modelled by one parameter, lateral displacement) hence the complexity of the fuzzy partitions is reduced [26].

The evolutionary process of CHC is illustrated in Figure 2. There are three key differences of CHC compared to a traditional GA. These differences are described here.

Cross-generational elitist selection

Selecting individuals for the next population occurs across parents and offsprings. This selection method is the same as that used in the $(\mu + \lambda)$ evolutionary strategy [34] where μ refers to the number of parents and λ indicates the number of offspring. Selection uses elitism to select the best μ parents from a population or the best λ offspring that are always copied over to the next population.

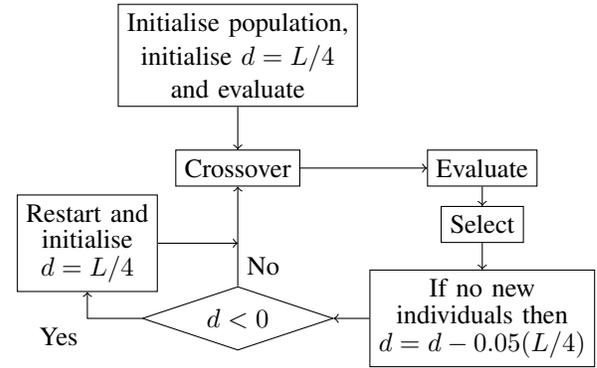


Fig. 2: Evolutionary process of CHC

Heterogeneous re-combination

Uniform crossover is applied where the probability of crossing over each bit in a binary representation (of L bits) is 50%, rather than crossing over segments of bits. Uniform crossover is said to be highly disruptive by [32] because it swaps about half the genes during crossover. An incest prevention mechanism prevents reproduction between similar chromosomes, so crossover is only performed on chromosomes whose measured difference is above a difference threshold d . Crossover is performed on two individuals if the Hamming distance between the chromosomes' bit strings is above the threshold.

Cataclysmic mutation

The mutation operator is not present in CHC. Instead, a restart operator provides the exploration ability that is crucial for a GA. Restart reintroduces diversity when the population converges and there have been no new chromosomes for multiple generations. Instead of mutating at every generation, a population is restarted in only those generations where the level of diversity drops below a threshold, which is determined by the incest prevention mechanism. Note that convergence is not used as a termination criterion. When a population is restarted, each individual is reinitialised except the best individual, which is just copied. Each individual is evaluated and the algorithm continues. A Boolean representation (of real value numbers) is used, and a percentage of bits is flipped. The percentage of bits is referred to as divergence rate [32]. The best individual is used as a template for creating other individuals.

The incest prevention mechanism slows convergence, which helps prevent premature convergence and is integral to CHC's operation because it influences crossover and restart.

Initialisation and the genetic operators are the same as [31]. In our proposal, we only change the fitness function.

A real-value representation of the 2-tuple linguistic representation of membership functions in a chromosome is defined as

$$\{c_{11}, \dots, c_{1m}, c_{21}, \dots, c_{2m}, \dots, c_{n1}, \dots, c_{nm}\},$$

where c is the lateral displacement, n is the number of items, and m is the number of linguistic labels. The lateral displacement for all linguistic terms of one item are followed by the lateral displacements for all linguistic terms of the next item and so on.

B. Fuzzy Transactions

The general model for fuzzy association rules [5] is used by the GA. In this general model, data is considered as a set of fuzzy transactions. A (crisp) set of fuzzy transactions is referred to as an FT-set. Consider an FT-set T and a set of items I . A fuzzy transaction is a non-empty fuzzy subset $\tilde{\tau} \subseteq I$. For every $i \in I$, there is a degree of membership $\tilde{\tau}(i)$ of item i in fuzzy transaction $\tilde{\tau}$. The degree of inclusion $\tilde{\tau}(J)$ of an itemset $J \subseteq I$ in fuzzy transaction $\tilde{\tau}$ is defined as

$$\tilde{\tau}(J) = \min_{i \in J} \tilde{\tau}(i). \quad (4)$$

For example [5], consider a set of items $I = \{i_1, i_2, i_3, i_4\}$ in a dataset of 6 transactions in Table I. Each row is a fuzzy transaction $\tilde{\tau}$, each column is an item i , and each cell value is a degree of membership $\tilde{\tau}(i)$. An example degree of inclusion is $\tilde{\tau}_1(\{i_3, i_4\}) = 0.7$. The dataset in Table I has an FT-set $\{\tilde{\tau}_1, \dots, \tilde{\tau}_6\}$, which is a crisp set.

An itemset J belongs to the fuzzy set $\tilde{\Gamma}_J$ on an FT-set according to

$$\tilde{\Gamma}_J(\tilde{\tau}) = \tilde{\tau}(J). \quad (5)$$

Scalar cardinality is calculated using the sigma count of members in the fuzzy set $\tilde{\Gamma}_J$.

TABLE I: Example set of fuzzy transactions $\tilde{\tau}$ containing items i and degrees of membership

	i_1	i_2	i_3	i_4
$\tilde{\tau}_1$	0	0.6	0.7	0.9
$\tilde{\tau}_2$	0	1	0	1
$\tilde{\tau}_3$	1	0.5	0.75	1
$\tilde{\tau}_4$	1	0	0.1	1
$\tilde{\tau}_5$	0.5	1	0	1
$\tilde{\tau}_6$	1	0	0.75	1

This general model for fuzzy association rules is applied only in the GA to measure scalar cardinality instead of fuzzy support.

C. Fitness Function

The fitness function of CHC from [35] was previously used to tune the membership functions by learning the lateral displacements [31]. In this paper, the fitness function was extended to tune larger membership grades using two proposed scalar cardinality measures.

An FT-set models every large 1-itemset I_k as a fuzzy set $\tilde{\Gamma}_{I_k}$ on the FT-set. The set of large 1-itemsets I is a

set of all 1-itemsets from a chromosome that are greater than or equal to the minimum fuzzy support threshold used in FuzzyApriori. Fuzzy support [36] [37] [11] of a large 1-itemset I_k is replaced with the scalar cardinality of the fuzzy set $\tilde{\Gamma}_{I_k}$.

The fitness value of a chromosome C_q is defined as

$$\text{fitness}(C_q) = \frac{\sum_{k=1}^n \text{cardinality}(\tilde{\Gamma}_{I_k})}{\text{suitability}(C_q)} \quad (6)$$

where I is the set of n large 1-itemsets that are determined by the membership functions in C_q . The cardinality measure replaces the fuzzy_support measure used in [35]. Two measures of scalar cardinality are proposed in the following subsections.

The suitability of membership functions in a chromosome C_q is measured by $\text{suitability}(C_q)$. Suitability measures the shape of the set of membership functions in a chromosome C_q and is defined as

$$\text{suitability}(C_q) = \sum_{k=1}^n \text{overlap_factor}(C_{qk}) + \text{coverage_factor}(C_{qk}), \quad (7)$$

where $\text{overlap_factor}(C_{qk})$ is the overlap factor of membership functions of item I_k in chromosome C_q , $\text{coverage_factor}(C_{qk})$ is the coverage factor of membership functions of item I_k in chromosome C_q , and n is the number of items in the chromosome. The aim of the overlap factor and the coverage factor is to enhance interpretability [35] [31]. The overlap factor prevents too much overlap of membership functions. Membership functions receive maximum penalty when they fully overlap. The coverage factor prevents separation between membership functions and promotes membership functions to cover the universe of discourse so that every data point belongs to a fuzzy set. The overlap factor and coverage factor address the semantic interpretability of fuzzy partitions [26].

The two proposed measures of scalar cardinality are:

1) α -cut thresholding: The α -cut of a fuzzy set A is defined [38] as

$$A_\alpha = \{x \mid \mu_A(x) \geq \alpha, \alpha \in [0, 1]\}, \quad (8)$$

which is a crisp set of elements of the domain with membership grades greater than or equal to the α level. A fuzzy set A can be represented by the union of its α -cut sets weighted by α [39] according to

$$A = \bigcup_{\alpha \in [0, 1]} \alpha \cdot A_\alpha. \quad (9)$$

This is referred to as the decomposition theorem that was proposed by Zadeh under the name resolution identity [40].

The first proposed scalar cardinality specifies the lower bound of the interval for the values of α . A fuzzy set A' is a

reduction of fuzzy set A represented by the union of α -cuts of A restricted by a lower bound a , and is defined as

$$A' = \bigcup_{\alpha \in [a, 1]} \alpha \cdot A_{\alpha}. \quad (10)$$

The set containing all the levels $\alpha \in [a, 1]$ of distinct α -cuts of a fuzzy set A' is the level set of A' , which is defined as

$$\Lambda_{A'} = \{\alpha \mid A'(x) = \alpha \text{ for some } x \in X, \alpha \geq a\}. \quad (11)$$

The α -cut thresholding scalar cardinality thresholds the membership grades for determining scalar cardinality of a fuzzy set.

2) *power-and-root α -cut weighting*: The second proposed scalar cardinality adjusts the α weight in the union operation. A piecewise function $\text{par}(\alpha)$ applies the n th root of α when $\alpha \geq a$, otherwise it raises the power of α to n . This is defined as

$$\text{par}(\alpha) = \begin{cases} \sqrt[n]{\alpha} & \text{if } \alpha \geq a; \\ \alpha^n & \text{otherwise.} \end{cases} \quad (12)$$

A fuzzy set A'' is represented by the union of its α -cut weighted by $\text{par}(\alpha)$ according to

$$A'' = \bigcup_{\alpha \in [0, 1]} \text{par}(\alpha) \cdot A_{\alpha}. \quad (13)$$

The power-and-root α -cut weighting scalar cardinality assigns more weight to large membership grades that are greater than or equal to a threshold a and less weight for small membership grades that are below the same threshold.

V. EXPERIMENTATION

The aim of experimentation was to identify if the two proposed measures of scalar cardinality can tune membership functions to have larger membership grades. The methodology, datasets, configuration of parameters, and results are presented.

A. Methodology

To identify if the fuzzy association rules have larger membership grades, a single run with/without the proposed scalar cardinality measures comprised the following steps:

- 1) Run GA to tune lateral displacements of membership functions.
- 2) Run FuzzyApriori to mine fuzzy association rules using the tuned membership functions.
- 3) Produce a normalised frequency distribution of every membership grade in the fuzzy set $\tilde{\Gamma}_{I_k}$ for an itemset I_k for all k rules found from FuzzyApriori.

The normalised frequency distribution had 20 bins for membership grades of width 0.05 in $[0.0, 1.0]$. A single run was performed with one of the proposed scalar cardinality measures and then repeated without that measure. The normalised frequency distribution found *without* a proposed scalar cardinality measure was subtracted from the normalised frequency

distribution found *with* a proposed scalar cardinality measure. The difference between the two normalised frequency distributions provides an indication of any increase or decrease in membership grades within interval widths of 0.05. Plotting the difference between two normalised frequency distributions allows qualitative and exploratory analysis of the proposed scalar cardinality measures.

B. Datasets

Four datasets from the Knowledge Extraction based on Evolutionary Learning (KEEL)¹ and University of California, Irvine Machine Learning (UCI)² data repositories were used. The datasets are listed in Table II with their properties. These datasets were chosen because they are real-world benchmark datasets and contain only real-value attributes.

TABLE II: Properties of the KEEL datasets

Name	Repository	# Instances	# Variables
Pollution	KEEL	60	16
Stock Price	KEEL	950	10
Stulong	KEEL	1419	5
Water Treatment Plant	UCI	537	38

C. Configuration of Parameters

CHC was configured with a population size of 50, 30 bits in a gene representation, 10,000 fitness evaluations, and the Parent Centric BLX Crossover was set to 1.0. Uniform fuzzy partitions for three linguistic terms were used for each variable.

The parameter n in the power-and-root α -cut weighting scalar cardinality was set to 2.

FuzzyApriori was run several times before experimentation to determine minimum fuzzy support and minimum confidence. For the purposes of demonstrating the idea proposed in this paper, the heuristic was to choose parameters that produced several hundred rules. The parameters for FuzzyApriori are shown in Table III.

TABLE III: FuzzyApriori parameters

Name	Minimum Fuzzy Support	Minimum Confidence
Pollution	0.6	0.8
Stock Price	0.35	0.8
Stulong	0.6	0.8
Water Treatment Plant	0.85	0.8

Implementations of CHC and FuzzyApriori from the KEEL software tool were used [41].

D. Results

A total of eight experiments were conducted. Four experiments made comparisons between using and not using the α -cut thresholding on the four datasets. Another four

¹<http://keel.es/>

²<http://archive.ics.uci.edu/ml/>

experiments made comparisons between using and not using the power-and-root α -cut weighting on the four datasets. In each experiment, the proposed scalar cardinality measures were repeated with different values of the a parameter, and the differences in normalised frequency distributions were produced.

Note, the minimum a value was different for each dataset, because the minimum support in FuzzyApriori was configured separately for each dataset. The fitness function only evaluates large 1-itemsets that meet the minimum support, and the minimum is used for the intersection in FuzzyApriori, so values of a below the minimum support threshold have no effect on tuning membership functions.

The difference in normalised frequency distributions of using α -cut thresholding and not using α -cut thresholding for various values of a on the Stulong dataset are shown in Figure 3a. The difference in normalised frequency distribution for $a = 0.9$ shows a decrease in the quantity of rules with membership grades between 0 and 0.8, and an increase in the quantity of rules with membership grades greater than 0.8. This shows that α -cut thresholding scalar cardinality is reducing the number of small membership grades and increasing the number of large membership grades. The same pattern is observed for values of a in $\{0.8, 0.85, 0.9\}$, however, this does not occur for other values of a . For the Pollution (Figure 3e) and Water Treatment Plant (Figure 3g) datasets, the same pattern is observed for values of a in $\{0.85, 0.9, 0.95\}$. A pattern illustrating the opposite effect is observed on the results for Stock Price dataset (Figure 3c) where all values of a show an increase in smaller membership grades and a decrease in larger membership grades. The α -cut thresholding measure was effective on three datasets for large values of a and not effective on the Stock Price dataset for all values of a .

The power-and-root α -cut weighting measure applied to the Stock Price dataset produced the normalised frequency distributions for various values of a shown in Figure 3d. This shows a decrease in smaller membership grades and an increase in larger membership grades for all values of a , which is the opposite to that observed on the same dataset with α -cut thresholding scalar cardinality. For all other datasets tested, the results of power-and-root α -cut weighting measure indicate no observable improvement, in terms of decreasing the number of small membership grades and increasing the number of large membership grades.

VI. CONCLUSIONS

Issues regarding the accuracy and interpretability of fuzzy association rules were discussed. Questions are raised about how accurately a fuzzy association rule with many small membership grades represents a relation in a dataset, and whether or not many small membership grades are useful in interpreting fuzzy association rules. Re-assessing the accuracy measures used in the tuning/learning of membership functions in FRBSs, fuzzy association rule mining, and more broadly fuzzy data mining can provide further insight into

the accuracy, interpretability, and issues around the trade-off between accuracy and interpretability.

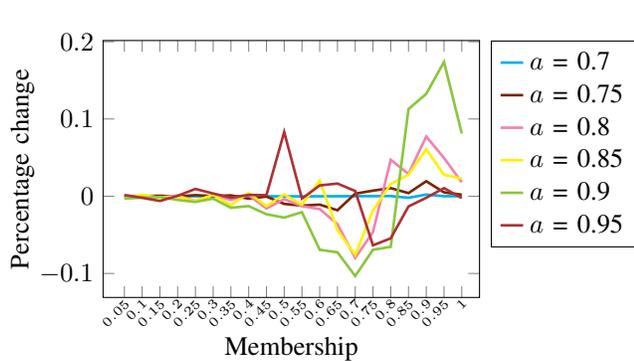
Two measures for scalar cardinality of a fuzzy set were proposed for tuning membership functions with a GA. The α -cut thresholding measure determines scalar cardinality of a fuzzy set by only including the union of α -cuts above a user-specified threshold. When applied to tuning membership functions, an increase in larger membership grades was observed on three of the four real-world benchmark datasets tested. The power-and-root α -cut weighting measure weights the α value used in the union of α -cuts, and an increase in larger membership grades was observed on the dataset that did not show the same increase with the α -cut thresholding measure. This suggests the choice of cardinality measure is associated with aspects or features of the dataset.

These preliminary results suggest the two proposed scalar cardinality measures are capable in guiding the tuning of membership functions with a GA towards larger membership grades on the datasets tested. Further analysis and more datasets are required to analyse why the scalar cardinality measures appear to be better suited to some datasets and not others. A complete comparative analysis of proposed scalar cardinality measures will be conducted.

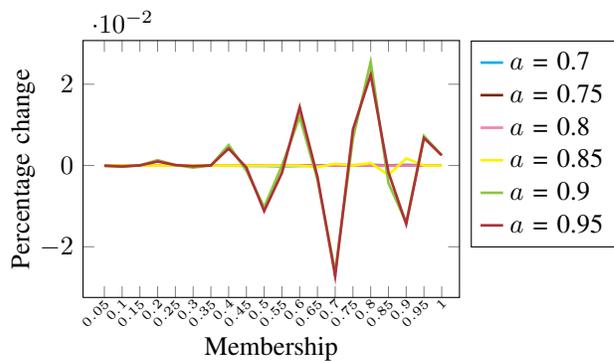
There is potential for further understanding how the learning and tuning of FRBSs and fuzzy association rules can address accuracy and interpretability problems associated with the sigma count problem.

REFERENCES

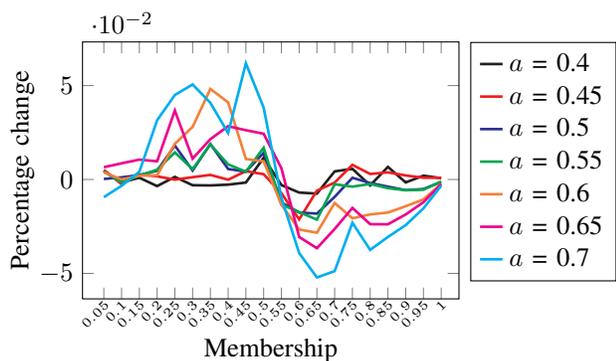
- [1] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," in *Proceedings of ACM SIGMOD International Conference on Management of Data, Washington, DC, USA*. New York: ACM, 1993, pp. 207–216.
- [2] R. Srikant and R. Agrawal, "Mining quantitative association rules in large relational tables," in *Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data*, ser. SIGMOD '96, 1996, pp. 1–12.
- [3] L. A. Zadeh, "Fuzzy sets," *Information Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [4] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in *Proceedings of the 20th International Conference on Very Large Data Bases*. San Francisco: Morgan Kaufmann Publishers Inc., 1994, pp. 487–499.
- [5] M. Delgado, N. Marín, D. Sánchez, and M.-A. Vila, "Fuzzy association rules: general model and applications," *Fuzzy Systems, IEEE Transactions on*, vol. 11, no. 2, pp. 214–225, 2003.
- [6] P. Bosc and O. Pivert, "On some fuzzy extensions of association rules," in *Joint 9th IFSA World Congress and 20th NAFIPS International Conference*, vol. 2, 2001, pp. 1104–1109.
- [7] D. Dubois, E. Hüllermeier, and H. Prade, "A systematic approach to the assessment of fuzzy association rules," *Data Mining and Knowledge Discovery*, vol. 13, no. 2, pp. 167–192, 2006.
- [8] R. R. Yager, "A new approach to the summarization of data," *Information Sciences*, vol. 28, no. 1, pp. 69–86, 1982.
- [9] A. Maeda, H. Ashida, Y. Taniguchi, and Y. Takahashi, "Data mining system using fuzzy rule induction," in *Proceedings of 1995 IEEE International Conference on Fuzzy Systems*, vol. 5, 1995, pp. 45–46.
- [10] K. C. C. Chan and W.-H. Au, "Mining fuzzy association rules," in *Proceedings of the Sixth International Conference on Information and Knowledge Management, Las Vegas, NV, USA, 1997*, pp. 209–215.
- [11] T.-P. Hong, C.-S. Kuo, and S.-C. Chi, "Trade-off between computation time and number of rules for fuzzy mining from quantitative data," *International Journal of Uncertainty and Fuzziness and Knowledge-Based Systems*, vol. 9, no. 5, pp. 587–604, 2001.



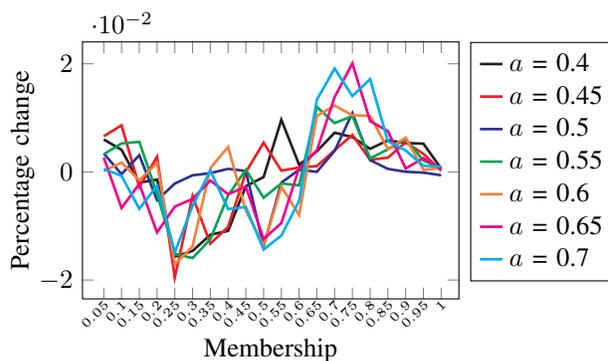
(a) Stulong dataset with α -cut thresholding



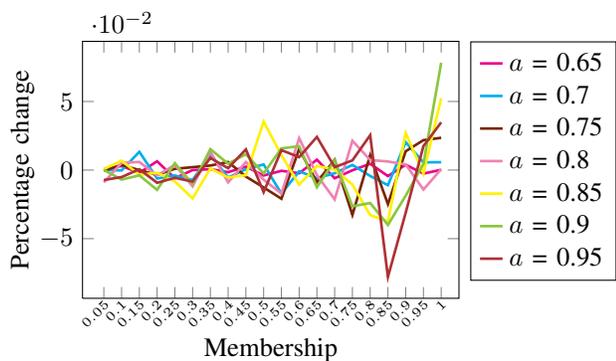
(b) Stulong dataset with power-and-root α -cut weighting



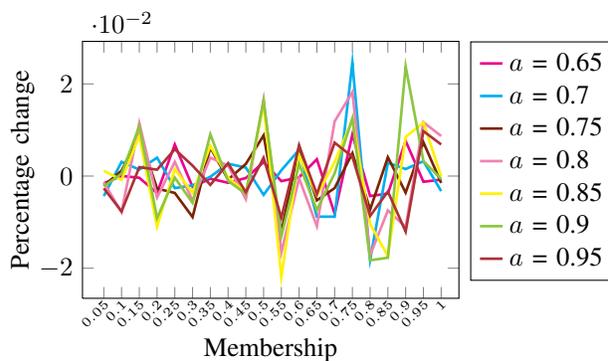
(c) Stock Price dataset with α -cut thresholding



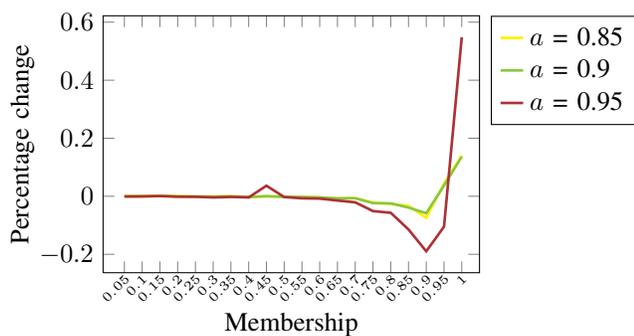
(d) Stock Price dataset with power-and-root α -cut weighting



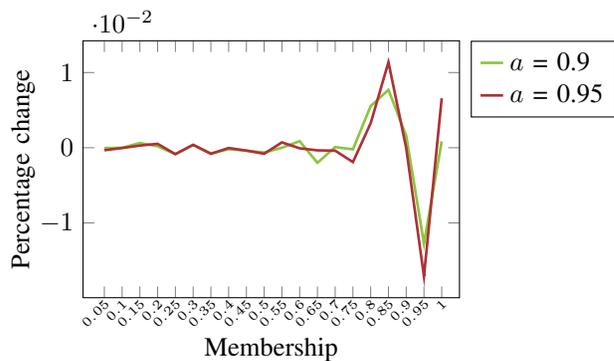
(e) Pollution dataset with α -cut thresholding



(f) Pollution dataset with power-and-root α -cut weighting



(g) Water Treatment Plant dataset with α -cut thresholding



(h) Water Treatment Plant dataset with power-and-root α -cut weighting

Fig. 3: Difference in normalised frequency distributions for all experiments

- [12] S. G. Matthews, M. A. Gongora, A. A. Hopgood, and S. Ahmadi, "Web usage mining with evolutionary extraction of temporal fuzzy association rules," *Knowledge-Based Systems*, vol. 54, no. 0, pp. 66–72, 2013.
- [13] W. H. B. Liu and Y. Ma, "Integrating classification and association rule mining," in *Proceedings of the International Conference on Knowledge Discovery and Data Mining*, 1998, pp. 80–86.
- [14] M. J. Martín-Bautista, D. Sánchez, M. A. V. Miranda, and H. L. Larsen, "Measuring effectiveness in fuzzy information retrieval," in *FQAS*, 2000, pp. 396–402.
- [15] A. D. Luca and S. Termini, "A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory," *Information and Control*, vol. 20, no. 4, pp. 301–312, 1972.
- [16] D. Ralescu, "Cardinality, quantifiers, and the aggregation of fuzzy criteria," *Fuzzy Sets and Systems*, vol. 69, no. 3, pp. 355–365, 1995.
- [17] M. Wygralak, "On the best scalar approximation of cardinality of a fuzzy set," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 5, no. 6, pp. 681–687, 1997.
- [18] D. Dubois and H. Prade, "Fuzzy cardinality and the modeling of imprecise quantification," *Fuzzy Sets and Systems*, vol. 16, no. 3, pp. 199–230, 1985.
- [19] M. Delgado, D. Sánchez, M. Martín-Bautista, and M. A. Vila, "A probabilistic definition of a nonconvex fuzzy cardinality," *Fuzzy Sets and Systems*, vol. 126, no. 2, pp. 177–190, 2002.
- [20] T. Martin and B. Azvine, "The X-mu approach: Fuzzy quantities, fuzzy arithmetic and fuzzy association rules," in *IEEE Symposium on Foundations of Computational Intelligence (FOCI)*, 2013, pp. 24–29.
- [21] D. Lewis and T. P. Martin, "X- μ fuzzy association rule method," in *Computational Intelligence (UKCI), 2013 13th UK Workshop on*, 2013, pp. 144–150.
- [22] D. Sánchez, M. Delgado, M. Vila, and J. Chamorro-Martínez, "On a non-nested level-based representation of fuzziness," *Fuzzy Sets and Systems*, vol. 192, no. 0, pp. 159–175, 2012.
- [23] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena, *Accuracy improvements in linguistic fuzzy modeling*, ser. Studies in fuzziness and soft computing. Berlin: Springer, 2003.
- [24] —, *Interpretability Issues in Fuzzy Modeling*, ser. Studies in fuzziness and soft computing. Berlin: Springer, 2003.
- [25] O. Cordón, "A historical review of evolutionary learning methods for mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems," *International Journal of Approximate Reasoning*, vol. 52, no. 6, pp. 894–913, 2011.
- [26] M. Gacto, R. Alcalá, and F. Herrera, "Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures," *Information Sciences*, vol. 181, no. 20, pp. 4340–4360, 2011.
- [27] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. SMC-15, no. 1, pp. 116–132, 1985.
- [28] M. Sugeno and G. Kang, "Structure identification of fuzzy model," *Fuzzy Sets and Systems*, vol. 28, no. 1, pp. 15–33, 1988.
- [29] J. H. Holland, *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI, 1975.
- [30] F. Herrera, "Genetic fuzzy systems: taxonomy, current research trends and prospects," *Evolutionary Intelligence*, vol. 1, no. 1, pp. 27–46, 2008.
- [31] J. Alcalá-Fdez, R. Alcalá, M. J. Gacto, and F. Herrera, "Learning the membership function contexts for mining fuzzy association rules by using genetic algorithms," *Fuzzy Sets and Systems*, vol. 160, no. 7, pp. 905–921, 2009.
- [32] L. J. Eshelman, "The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination," in *Foundations of Genetic Algorithms*, vol. 1. Morgan Kaufmann, 1991, pp. 265–283.
- [33] F. Herrera and L. Martínez, "A 2-tuple fuzzy linguistic representation model for computing with words," *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 6, pp. 746–752, 2000.
- [34] H.-P. Schwefel, "Evolution strategy and numerical optimization," Ph.D. dissertation, Technical University of Berlin, 1975.
- [35] T.-P. Hong, C.-H. Chen, Y.-C. Lee, and Y.-L. Wu, "Genetic-fuzzy data mining with divide-and-conquer strategy," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 2, pp. 252–265, 2008.
- [36] R. R. Yager, "Database discovery using fuzzy sets," *International Journal of Intelligent Systems*, vol. 11, no. 9, pp. 691–712, 1996.
- [37] H. Ishibuchi, T. Nakashima, and T. Yamamoto, "Fuzzy association rules for handling continuous attributes," in *Proceedings of IEEE International Symposium on Industrial Electronics 2001*, vol. 1, 2001, pp. 118–121.
- [38] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall, 1997.
- [39] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning," *Information sciences*, vol. 8, no. 3, pp. 199–249, 1975.
- [40] L. Zadeh, "Similarity relations and fuzzy orderings," *Information Sciences*, vol. 3, no. 2, pp. 177–200, 1971.
- [41] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez, and F. Herrera, "KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework," *Journal of Multiple-Valued Logic and Soft Computing*, vol. 17, no. 2–3, pp. 255–287, 2011.