A Framework for Mining Fuzzy Association Rules from Composite items

M. Sulaiman Khan[1], Maybin Muyeba[1], Frans Coenen[2]
[1] School of Computing, Liverpool Hope University, UK,
[2] Department of Computer Science, University of Liverpool, UK,
{khanm@hope.ac.uk, muyebam@hope.ac.uk, frans@csc.liv.ac.uk}

Abstract. A novel framework is described for mining fuzzy Association Rules (ARs) relating the properties of composite attributes, i.e. attributes or items that each feature a number of values derived from a common schema. To apply fuzzy Association Rule Mining (ARM) we partition the property values into fuzzy property sets. This paper describes: (i) the process of deriving the fuzzy sets (Composite Fuzzy ARM or CFARM) and (ii) a unique property ARM algorithm founded on the correlation factor interestingness measure. The paper includes a complete analysis, demonstrating: (i) the potential of fuzzy property ARs, and (ii) that a more succinct set of property ARs (than that generated using a non-fuzzy method) can be produced using the proposed approach.

Keywords: Association rules, fuzzy association rules, composite attributes, quantitative attributes.

1 Introduction

Association Rule Mining (ARM) is an important and well established data mining topic. The objective of ARM is to identify patterns expressed in the form of Association Rules (ARs) in transaction data sets [5, 6, 12]. The attributes in ARM data sets are usually binary valued but ARM has also been applied to quantitative and categorical (non-binary) data [1, 13, 16]. With the latter, values can be split into linguistically labeled ranges (for example “low”, “medium”, “high” etc) such that each range represents a binary valued attribute. Values can be assigned to these range attributes using crisp or fuzzy boundaries. The application of ARM using the latter is referred to as fuzzy ARM (FARM) [1]. Fuzzy ARM has been shown to produce more expressive ARs than the “crisp” methods [1, 3, 4,].

In this paper we introduce the problem of “Composite item” Fuzzy ARM (CFARM) whose main objective is the generation of fuzzy ARs associating the “properties” linked with composite attributes [15] i.e. attributes or items composed of sets of sub-attributes that conform to a common schema. An example could be market basket analysis where the attribute set I is a set of groceries, and P is a set of nutritional properties (P) that these groceries posses (i.e. P = {Pr, Fe, Ca, Cu,..}) standing for protein, Iron etc. Note that the actual values (properties) associated with each element of I will be constant.

The main contributions of this paper are: (i) a framework for the "Composite item" mining of property ARs and (ii) an evaluation of the potential of using property ARs. In particular we demonstrate that the proposed approach generates a more succinct set of property ARs (than that generated using a non-fuzzy method).

The paper is organised as follows. In section 2 we present the background and related work for the proposed composite fuzzy ARM approach. Section 3 presents a sequence of formal definitions for the work; and section 4 details the CFARM algorithm; section 5 expands the motivation with an example application; section 6 gives a complete analysis of the CFARM algorithm, and section 7 concludes the paper with a summary of the contribution of the work and directions for future work.

2 Background and Related Work

The term composite item has been used previously in the context of data mining. In [8, 16], a composite item is defined as a combination of several items, e.g. if itemset {A, B} and {A, C} are not large then rules {B}→{A} and {C}→{A} will not be generated, but by combining B and C to make a new *composite* item {BC} which may be large, rules such as {BC}→{A} may be generated. In this paper we define composite items differently, as indicated earlier, to be an item with properties (a formal definition is presented in Section 3). The definition concurs with [15], the earliest references to composite attributes (that the authors are aware of). ARM usually uses binary valued attributes, quantitative attributes usually discretised into partitions resulting in the "sharp boundary" problem. Fuzzy ARM [3, 7, 14] has been shown to resolve this problem.

To illustrate the concept of Fuzzy ARM applied to composite ARM, we consider super market basket analysis where the set of groceries (I) have a common set of nutritional quantitative properties (Table 1).

Table 1. Example composite attributes (*groceries*) with their associated properties (*nutrients*)

Items/Nutrients	Protein	Fibre	Carbohydrate	Fat	…
Milk	3.1	0	4.7	.2	…
Bread	8	3.3	43.7	1.5	…
Biscuit	6.8	4.8	66.3	22.8	…
…	…	…	…	…	…

3 Problem Definition

In this section formal definitions are presented to define composite attributes, the FARM concept and the normalization process for Fuzzy Transactions (*FT)* .

3.1 Formal definitions

Definition 1: Fuzzy Association Rules

A Fuzzy AR [3] is an implication of the form: if $\langle A, X \rangle$ then $\langle B, Y \rangle$, where A and B are disjoint itemsets and X and Y are fuzzy sets. In our case the itemsets are made up of property attributes and the fuzzy sets are identified by linguistic labels.

Definition 2: Raw Dataset

A Raw Dataset D consists of a set of transactions $T = \{t_1, t_2, t_3,..,t_n\}$, a set of composite items $I = \{i_1, i_2, i_3,.., i_{|I|}\}$ and a set of properties $P = \{p_1, p_2, p_3,.., p_m\}$. Thus each item i_j will have associated with it a set of values corresponding to the set P, i.e. $t_i[i_j] = \{v \mid v_1, v_2, v_3,.., v_m\}$. The “kth” property value for the “jth” item in the “ith” transaction is given by $t_i[i_j[v_k]]$.

Note that a property attribute can take either a categorical or a quantitative value and denoted as <label,value> (see Table 2). In the rest of this paper the term “item” means an *item* in an itemset in the manner associated with traditional ARM, and the term *attribute* is used to mean a property item (sub-item).

Table 2. Example raw dataset *D*

TID	Record
1	{<a,{2,4,6}>, <b,{4,5,3}>}
2	{<c,{1,2,5}>, <d,{4,2,3}>}
3	{<a,{2,4,6}>, <c,{1,2,5}>, <d,{4,1,3}>}
4	{<b,{4,5,3}>, <d,{4,2,3}>}

Definition 3: Property Dataset

In the process described here, the given raw dataset D is initially transformed into a property data set D^P which consists of property transactions $T^P = \{t_1^P, t_2^p, t_3^p,.., t_n^p\}$ and a set of property attributes P (instead of a set of composite items I). Each transaction t_i^P (the “ith” transaction) is some subset of $P = \{p_1, p_2, p_3,.., p_m\}$. The value for each property attribute $t_i^P[p_j]$ (the “jth” property attribute in the “ith” property transaction) has a value obtained by aggregating the numeric values for all p_j in t_i (See Table 3). Thus equation 1:

$$\text{Prop Value}(t_i^p[p_j]) = \frac{\sum_{j=1}^{|t_i|} t_i[i_j[v_k]]}{|t_i|} \quad \textbf{(1)}$$

Table 3. Example property data set D^p generated raw data set given in table 2

TID	X	Y	Z
1	3.0	4.5	4.5
2	3.0	2.0	4.0
3	2.3	2.3	4.7
4	4.0	3.5	3.0

Definition 4: Fuzzy Dataset

Once a property data set D^p is defined, it is then transformed into a fuzzy dataset D'. A fuzzy dataset D' consists of fuzzy transactions $T' = \{t'_1, t'_2, t'_3, ..., t'_n\}$ and a set of fuzzy property attributes P' each of which has fuzzy sets with linguistic *labels* $L = \{l_1, l_2, l_3, ..., l_{|L|}\}$. Each property attribute $t_i^p[p_j]$ is associated (to some degree) with several fuzzy sets and given by a *membership degree* value, in the range $[0..1]$, which indicates the correspondence between the value of a given $t_i^p[p_j]$ and the set of *fuzzy linguistic labels*. The "kth" label for the "jth" property attribute for the "ith" fuzzy transaction is given by $t'_i[p_j[l_k]]$.

The nature of the user defined fuzzy ranges is expressed in a *properties table* (see definition 6 below). The numeric values for each property attribute $t_i^p[p_j]$ are *fuzzified* (mapped) into the appropriate membership degree values using a membership function $\mu(t_i^p[p_j], l_k)$ that applies the value of $t_i^p[p_j]$ to a label $l_k \in L$, e.g.

$$t'_i[p_j] = \{\mu(t_i^p[p_j]], l_1), \mu(t_i^p[p_j]], l_2), \mu(t_i^p[p_j]], l_3), .., \mu(t_i^p[p_j]], l_{|L|})\}$$

The nature of the function is discussed in more detail in Sub-section 3.2 below. The complete set of fuzzy property attributes P' is then given by $P \times L$. A fuzzy data (Table 4) based on the property data set (Table 3) are given. The membership values are all normalised to contribute support counts of 0 or 1 for a single attribute in a single record .

Table 4. Example Fuzzy data set ($L = \{\text{small}, \text{medium}, \text{large}\}$, μ unspecified).

TID	X			Y			Z		
	Small	Medium	Large	Small	Medium	Large	Small	Medium	Large
1	0.5	0.5	0.0	0.0	0.5	0.5	0.0	0.0	1.0
2	0.0	0.0	1.0	0.0	0.0	1.0	1.0	0.0	0.0
3	1.0	0.0	0.0	1.0	0.0	0.0	0.0	0.5	0.5
4	0.0	0.0	1.0	1.0	0.0	0.0	1.0	0.0	0.0

Definition 5: Composite Itemset Value Table.

A Composite Itemset Value (CIV) table allows ac cess to property values for specific items. Note that a CIV table is not always required; the values may be included in the raw data as in the case of the example raw dataset presented in Table 2 where property values are all in the range [1..6]. In some applications specific attributes always have the same property values, which is the case in the extended market basket analysis example, introduced in Section 1.

For completeness, the CIV table for the example raw dataset given in Table 2 is given in Table 5 below.

Table 5. Composite Itemset Value Table for raw dataset given in Table 2

Item	Property attributes		
	X	Y	Z
A	2	4	6
B	4	5	3
C	1	2	5
D	4	1	3

Table 6. Property Table for raw dataset given in Table 2

Property	Linguistic values		
	Low	Medium	High
X	$v_k \leq 2.3$	$2.3 < v_k \leq 3.7$	$3.7 < v_k$
Y	$v_k \leq 3.3$	$3.3 < v_k \leq 4.7$	$4.7 < v_k$
Z	$v_k \leq 4.3$	$4.3 < v_k \leq 5.7$	$5.7 < v_k$

Definition 6: Properties Table

A Properties Table maps all possible values for each property attribute $t_i^p[p_j]$ onto user defined (overlapping) ranges, each associated with a linguistic label from labels in L. Property tables provide a mapping of property attribute values to membership values. An example is given in Table 6 for the raw data set (Table 2).

Definition 7: Fuzzy Frequent Itemsets
A property attribute set A, $A \subseteq P \times L$, is a fuzzy frequent attribute set if its *fuzzy support value* is greater than or equal to a minimum support threshold (the notion of fuzzy support values is discussed further in sub-section 3.3 below). The significance is that fuzzy ARs are generated from discovered frequent attribute sets.

Definition 8: Fuzzy Normalisation
Fuzzy normalisation is the process of finding the contribution to the fuzzy support value, m', for individual property attributes ($t_i^p[p_j[l_k]]$) such that a partition of unity is guaranteed. This is given by equation 2 (μ is the membership function). Without normalisation, the sum of the support contributions of individual fuzzy sets associated with an attribute in a single transaction may no longer be unity.

$$t_i'[p_j[l_k]] = \frac{\mu(t_i^p[p_j[l_k]])}{\sum_{x=1}^{|L|} \mu(t_i^p[p_j[l_x]])} \qquad \textbf{(2)}$$

3.2 Membership Function

The membership degree to a particular fuzzy set, $t_i[p_j[l_k]]$ is determined by a membership function, of which there are many different types. An example, using the market basket analysis introduced earlier, is given in Fig 1 with the membership functions for the Protein nutrient.

With respect to the application, the trapezoidal shape was chosen as it best captured the intuition (promoted by nutritionists) that nutrient values above or below the ideal value 1 are undesirable. A function representing all the membership degrees of an input value "x" has the letters α, β, γ and δ to refer to the corners of the trapezium proceeding in a clockwise fashion starting with the bottom-left corner. The value x has an "ideal" value between the points β to γ along the "X" axis, with the lowest value α and the highest value δ. For missing values or so called "trace" elements, the fuzzy function evaluates to zero degree membership.

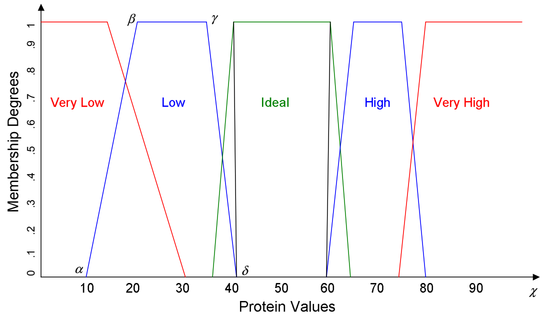

Fig. 1. Fuzzy Membership functions

3.3 Fuzzy Support and Confidence

The support-confidence framework remains the most popular in traditional ARM. With some adjustment the support-confidence framework can be applied to fuzzy ARM. Frequent fuzzy attribute sets are identified by calculating fuzzy support (significance) values. Fuzzy Support (FS) is typically calculated as follows [1]:

$$FS(A) = \frac{\text{Sum of votes satisfying A}}{\text{Number of records in } T} \tag{3}$$

where $A = \{a_1, a_2, a_3, \ldots, a_{|A|}\}$ is a set of property attribute-fuzzy set (label) pairs such that $A \subseteq P \times L$. A record t_i' "satisfies" A if $A \subseteq t_i'$. The individual vote per record t_i, is obtaining by multiplying the membership degree associated with each attribute-fuzzy set pair $[i[l]] \in A$:

$$\text{vote for } t_i \text{ satisfying } A = \prod_{\forall [i[l]] \in A} t_i'[i[l]] \tag{4}$$

So we have,

$$\text{FS}(A) = \frac{\sum_{i=1}^{i=n} \prod_{\forall [i[l]] \in A} t_i'[i[l]]}{n} \tag{5}$$

Note that by using the product operator (often referred to in fuzzy ARM literature as the *mul* operator) for fuzzy aggregation, the degree of contribution of all items is taken into account and thus provides for a more effective result.

Frequent attribute sets with fuzzy support above the user specified threshold are used to generate all possible rules. A fuzzy AR derived from a fuzzy frequent attribute set is of the form:

$$A \rightarrow B$$

where A and B are disjoint subsets of the set $P \times L$ such that $A \cup B = C$. Fuzzy Confidence (FC) (or fuzzy *certainty factor*) is calculated in the usual manner:

$$FC(A \to B) = \frac{FS(A \cup B)}{FS(A)} \tag{6}$$

3.4 Fuzzy Correlation

The Fuzzy Confidence measure (FC) described above does not use $FS(B)$, the fuzzy correlation measure (FCORR) addresses this. The correlation measure is a statistical measure founded on the concepts of *covariance* (Cov) and *variance* (Var) :

$$FCORR(A \to B) = \frac{Cov(A,B)}{\sqrt{Var(A) \times Vat(B)}} \tag{7}$$

The value of correlation ranges from -1 to +1. Value -1 means no correlation and +1 means maximum correlation. Thus we are only interested in rules that have a correlation value that is greater than 0. As the certainty value increases from 0 to 1, the more related the attributes are and consequently the more interesting the rule.

4 The CFARM Algorithm

For fuzzy ARM standard algorithms can be used or at least adapted after some modifications [12]. Less attention has been given to developing dedicated efficient algorithms for fuzzy ARM [5] but still there are some contributions in this area [7]. An efficient algorithm is required because a significant amount of processing (filtration, conversions, normalization) is undertaken to prepare the raw data prior to the application of fuzzy ARM.

The proposed Composite Fuzzy ARM (CFARM) algorithm belongs to the *breadth first traversal* family of ARM algorithms and works in a fashion similar to the Apriori algorithm [5]. The CFARM algorithm consists of four major steps:

1. Transformation of ordinary transactional data set (T) into a property data set (T^p).
2. Transformation of property data set (T^p) into a fuzzy data set T'.
3. Apply an Apriori style fuzzy ARM algorithm to T' using fuzzy support, confidence and correlation measures of the form described above to produce a set of frequent item sets F .
4. Process F and generate a set of fuzzy ARs R such that $\forall r \in R$ the certainty factor (either confidence or correlation as desired by the end user) is above some user specified threshold.

The algorithms for steps 1 and 2 is not shown here but we show examples of its application using a fragment of a the raw data set (T) given in Table 8(a). This raw data is then cast into a properties data set (T^P). This is done, as described above, by averaging the property values for each transaction (see definition 3 and table 4). The result is as shown in Table 8(b) which is then cast into a fuzzy data set T' as shown in Table 8(c). An alternative approach is to discretise the data.

Table 7. Some conventional datasets (*raw, property and conventional*)

(a) Raw data (T)

TID	Items
1	a, b
2	c
3	a, b, d
4	…

(b) Property data set (T^P)

TID	X	Y	Z
1	3.0	4.5	4.5
2	1	2	5
3	3.3	3.3	4.0
4	…	…	…

(c) *Fuzzy data set* (T')

TID	X			Y			Z		
	S	M	L	S	M	L	S	M	L
1	0.0	0.5	0.5	0.0	0.0	1.0	0.0	0.0	1.0
2	1.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0
3	0.0	0.2	0.8	0.2	0.8	0.0	0.5	0.5	0.0
4	…	…	…	…	…	…	…	…	…

5 An Example Application

To evaluate our approach, we used a market basket analysis data set with 600 composite edible items; the objective is to determine consumers' consumption patterns for different nutrients using RDA. The properties for each item comprised the 27 nutrients contained in the government sponsored RDA table (a partial list consists of Biotin, Calcium, Carbohydrate, .., Vitamin K, Zinc). These RDA values represent a CIV table. The property data set will therefore comprise $600 \times 27 = 16200$ attributes. The linguistic label set L was defined as {Very Low (VL), Low (L), Ideal (I), High (H), Very High (VH)}. Thus the set of fuzzy attributes $A = PXL$ has $27 \times 5 = 135$ attributes. A fragment of this data is given in Table 10.

A representative fragment of a raw data set (T), comprising edible items, is given in Table 11(a). This raw data is then cast into a properties data set (T^P) using the given CIV/RDA table to give the properties data set in Table 11(b). It is feasible to have alternative solutions here but we choose to code fuzzy sets {very low, low, ideal, high, very high} with numbers {1, 2, 3, 4, 5} for the first nutrient (Pr), {6, 7, 8, 9, 10} for the second nutrient (Fe) etc [10]. Thus, data in Table 11(c) can be used by any binary ARM algorithm.

Table 8. Fragment of market basket composite item data set[1].

Nutrients/Fuzzy Ranges	Very Low				Low				Ideal				High				Very High		
	Min	Core		Max	Min	Core		Max	Min	Core		Max	Min	Core		Max	Min	Core	
Fiber	0	1	10	15	10	15	20	25	20	25	30	35	30	33	38	39	35	40	…
Iron	0	.6	8	12	8	12	16	18	16	18	19	20	19	20	22	23	22	23	…
Protein	0	1	15	30	10	20	35	40	35	40	60	65	60	65	75	80	75	70	…
VitaminA	0	15	150	200	150	200	300	400	300	350	440	500	440	490	550	600	550	600	…
Zinc	0	.8	8	10	8	10	15	20	15	20	30	40	30	40	46	50	46	50	…
	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…

Table 9. Example data fragment from example application

(a) Raw data (T)

TID	Items
1	X, Z
2	Z
3	X,Y, Z
4	…

(b) Property data set (T^P)

TID	Pr	Fe	Ca	Cu
1	45	150	86	28
2	9	0	47	1.5
3	54	150	133	29.5
4	…	…	…	…

(c) Conventional ARM data set

TID	Pr	Fe	Ca	Cu
1	3	8	13	16
2	1	6	12	16
3	3	8	15	16
4	…	…	…	...

This approach only gives us the total support of various fuzzy sets per nutrient and not the degree of (fuzzy) support. This directly affects the number and quality of rules (see section 6). To tackle this, the fuzzy approach here converts the RDA property data set (Table 11(b)) to linguistic values (Table 12) for each nutrient and their corresponding degrees of membership reflected in each transaction. Table 12 shows only two nutrients, Pr and Fe (i.e. a total of 10 fuzzy sets). The CFARM algorithm uses a tree data structure to store itemsets.

Table 10. Linguistic transaction file

TID	Protein (Pr)					Iron (Fe)					
	VL	L	Ideal	H	VH	VL	L	Ideal	H	VH	…
1	0.0	0.7	0.3	0.0	0.0	0.0	0.0	0.8	0.2	0.0	…
2	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	…
3	0.0	0.0	0.9	0.1	0.0	0.0	0.0	0.8	0.2	0.0	…
4	…	…	…	…	…	…	…	…	…	…	…

[1] *Values could be in grams, milligrams, micrograms, International unit or any unit).Here Min is the minimum value i.e.* α *, Core is the core region* β, δ *and Max is the maximum value* γ *in the fuzzy membership graph.of figure 2.*

6 Experimental Results

To demonstrate the effectiveness of the approach, we performed several experiments using a T10I4N0.6KD100k (average of 10 items per transaction, average of 4 items per interesting set, 600 attributes and 100,000 transactions/records) QUEST data set [11]. Each of the 600 attributes was matched to one of 600 food items listed in a real RDA table. The data is thus a transactional database containing 100K records.

Our experiments in the first instance compared CFARM, with and without normalisation, against standard (discrete) ARM with respect to: (i) the number of frequent sets generated and (ii) the number of rules generated (using both the confidence and the correlation measure). Fig 2 shows the results and demonstrates the difference between the number of frequent itemsets generated using:

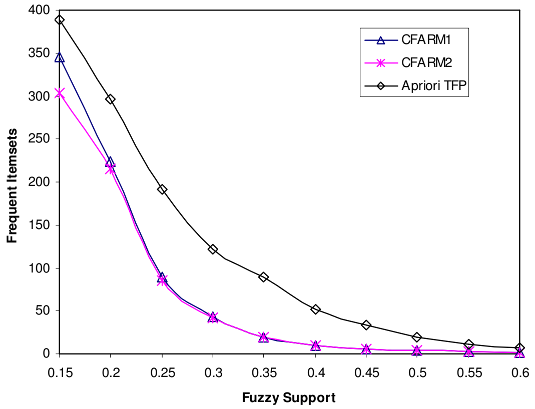

Fig. 2. Number of frequent Itemsets

1. Standard ARM using discrete intervals,
2. CFARM with fuzzy partitions without normalization (CFARM1), and
3. Fuzzy ARM with fuzzy partitions with normalization (CFARM2).

For standard ARM, the Apriori-TFP algorithm was used [6] with a range of support thresholds. As expected the number of frequent itemsets increases as the minimum support decreases.

In Fig 2, CFARM1 uses a dataset without normalization while CFARM2 uses dataset with normalization. From the results, it is clear that standard ARM produces more frequent itemsets (and consequently rules) than fuzzy ARM. This is because the frequent itemsets generated more accurately reflect the true patterns in the data set than the numerous artificial patterns resulting from the use of crisp boundaries in standard ARM. At low support threshold levels, the approach with normalization (CFARM2) starts to produce less frequent itemsets than the approach without normalization (CFARM1). This is because the average contribution to support counts per transaction is greater without using normalization than with normalization. Figs 3 and 4 compares number of rules generated using user specified fuzzy confidence and fuzzy correlation values respectively. In both cases, the number of rules is less using CFARM2; this is a direct consequence of the fact that CFARM 2 generates fewer frequent itemsets.

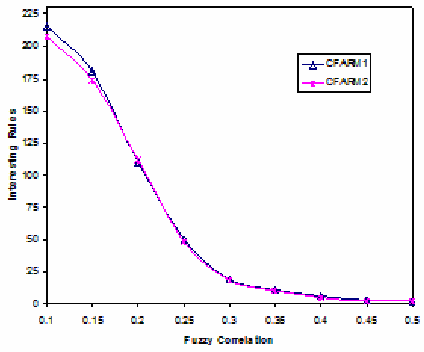

Fig. 3. No. of Rules (confidence)

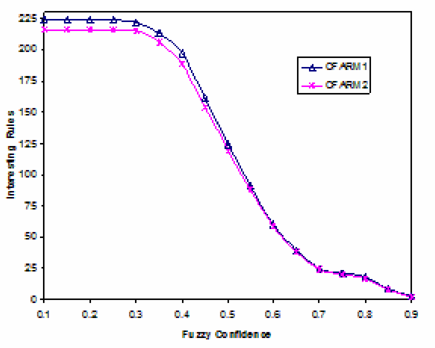

Fig. 4. No. of Rules (Correlation)

Note that fewer, but arguably better (succinct) rules are generated using the correlation measure (Fig 3) than the confidence measure (Fig 4). The experiments show that using the proposed fuzzy normalization process less fuzzy ARs are generated.

In general, we can see that the novelty of the approach is the ability to analyse datasets that can be expressed as composite items where each item has a number of property values. In addition, the approach shows that a more succinct set of property ARs than that generated using a non-fuzzy method) can be produced.

7. Conclusion and future work

In this paper, we have presented a novel framework for extracting hidden information from composite items where such have common properties defined as quantitative (sub) itemsets. The properties are then transformed into fuzzy sets. The CFARM algorithm produces a more succinct set of fuzzy ARs using fuzzy measures and correlation as the interestingness (certainty) measure and thus presents a new way for extracting ARs from items with properties. This is different from normal quantitative ARM. We also showed a practical example with market basket data where edible items were used with nutritional content as properties. Of note is the significant potential to apply CFARM to other applications where items could have composite attributes even with varying fuzzy sets between attributes. Overall, the approach presented here is effective for analysing databases with composite items. Further work will compare performance of the CFARM algorithm with common fuzzy ARM algorithms [1, 3].

References

1. Gyenesei, A.: A Fuzzy Approach for Mining Quantitative Association Rules, Acta Cybernetical, Vol. 15, (2), 305--320 (2001)
2. Lee, C. H., Chen, M. S., Lin, C. R.: Progressive Partition Miner, an Efficient Algorithm for Mining General Temporal Association Rules, IEEE Trans. on Knowledge and Data Engineering, Vol. 15, (4) pp. 1004-1017 (2003)
3. Kuok, C., Fu, A., Wong, H.: Mining Fuzzy Association Rules in Databases, ACM SIGMOD Record Vol. 27, (1), pp. 41-46 (1998)
4. Dubois, D., Hüllermeier, E., Prade, H.: A Systematic Approach to the Assessment of Fuzzy Association Rules, Data Mining and Knowledge Discovery Journal, Vol. 13(2), pp. 167-192 (2006)
5. Bodon, F.: A Fast Apriori implementation, in: Proc. (FIMI'03), IEEE ICDM Workshop on Frequent Itemset Mining Implementations, Vol. 90, Florida, USA, (2003)
6. Coenen, F., Leng, P., Goulbourne, G.: Tree Structures for Mining Association Rules, Data Mining and Knowledge Discovery, Vol. 8, (1), pp. 25 – 51 (2004)
7. Chen, G., Wei, Q.: Fuzzy Association Rules and the Extended Mining Algorithms, Information Sciences, Vol. 147, (1-4) pp. 201–28 (2002)
8. Wang, K., Liu, J. K., Ma, W.: Mining the Most Reliable Association Rules with Composite Items, in: Proc. ICDMW'06, pp. 749-754 (2006)
9. Delgado, M., Marin, N., Sanchez, D., Vila, M.A.: Fuzzy Association Rules, General Model and Applications, IEEE Transactions on Fuzzy Systems, 11(2) 214–225 (2003)
10. Muyeba, M., Sulaiman, M., Malik, Z., Tjortjis, C.: Towards Healthy Association Rule Mining (HARM), A Fuzzy Quantitative Approach, Proc. IDEAL'06, LNCS, Vol. 4224, 1014--1022 (2006)
11. Agrawal, R., Srikant, R.: Quest Synthetic Data Generator. IBM Almaden Research Center.
12. Agrawal, R., Imielinski, T., Swami, A.: Mining Association Rules Between Sets of Items in Large Databases, Proc. ACM SIGMOD Int. Conf. on Management of Data, Washington, D.C., pp. 207-216 (1993)
13. Srikant, R., Agrawal, R.: Mining Quantitative Association Rules in Large Relational Tables, Proc. ACM SIGMOD Conf. on Management of Data. ACM Press, Montreal, Quebec, pp. 1-12, (1996)
14. Au, W. H., Chan, K.: Farm, A Data Mining System for Discovering Fuzzy Association Rules, Proc. 8th IEEE Int'l Conf. on Fuzzy Systems, Seoul, Korea, pp. 1217-1222, (1999)
15. Kim, W., Bertino, E., Garza, J.: Composite objects revisited, ACM SIGMOD Record, Vol. 18, (2) (1989) 337-347
16. Ye, X., Keane, J.A: Mining Composite Items in Association Rules, Proc. IEEE Int. Conf. on Systems, Man and Cybernetics, pp. 1367--1372 (1997)