New Approach to Mining Fuzzy Association Rule with Linguistic Threshold Based on Hedge Algebras

Le Anh Phuong[1], Tran Dinh Khang[2], Nguyen Vinh Trung[3]

[1] Department of Computer Science, Hue University of Education, Hue University
[2] SoICT, Hanoi University of Science and Technology
[3] Information Technology Center, Hue University of Education, Hue University

Abstract. The authors [2-5] have studied and presented the quantitative method of linguistic variables and linguistic threshold by fuzzy set. Chien-Hua Wang, Chin-Pang Tzong proposed an algorithms for mining fuzzy association rule [2]. In this paper, we extend the algorithms proposed in [2] for number data and linguistic variables by using hedge algebras.

Keywords: fuzzy association rules, linguistic threshold, hedge algebra

1 Introduction

Data mining with the approach of association rules is one of important aspects in the field of data mining.

Many authors have presented various methods, algorithms of data mining by association rules with numerical support and confidence value. However, in reality, these values are natural linguistic ones. Besides, importance value of each item is evaluated not only by quantity, frequency of occurrence in each transaction but also by the qualitative evaluation of administrators (for those items) by natural language. And hedge algebra have met the requirements for directly processing calculation on linguistic value (without fuzzification, but with direct calculation based on qualitative semantic function and flexible calculation). Thus, it is necessary to establish a method of data mining by association rules with hedge algebra, in which the input is qualitative transactional database and qualitative evaluation table of those database items and the support, confidence values are also natural language ones.

2 Knowledge Base

2.1 Association rules

Let $I = I_1, I_2, \ldots, I_m$ be a set of items. Let D, the task-relevant data, be a set of database transactions where each transaction T is a set of items, such is $T \subseteq I$. Each transaction is associated with an identi_er, called TID.

Definition 1. *An association rule has the form of $X \rightarrow Y$, where $X \subseteq I$, $Y \subseteq I$, and $X \cap Y = \theta$.*

Definition 2. *The support of association rule $X \rightarrow Y$ the probability that $X \cup Y$ exists in a transaction in the database D.*

$$support(X \rightarrow Y) = \frac{|X \cap Y|}{|N|}$$

Definition 3. *The confidence of the association rule $X \rightarrow Y$ is the probability that $X \cup Y$ exists given that a transaction contains X, i.e.*

$$confidence\,(X \rightarrow Y) = \frac{support(X \cup Y)}{support(X)} = \frac{|X \cap Y|}{|X|}$$

Where: $|X|$ is the number of transactions, including X; $|X \cap Y|$ is the number of transactions, including X and Y; N is the total of transaction database.

Mining the association rules of the database is finding all of the rules that have the degree of support and confidence greater than degree of support minsup and confidence minconf determined by the available user.

2.2 Hedge algebras (HA)

Let **X** be a linguistic variable and **X** be a set of its terms, called a term-domain of **X**. E.g. if **X** is the rotation speed of an electrical motor and linguistic hedges used to describe its speed are $Very$, $More$, $Possibly$, $Little$, denoted correspondingly for short by **V**, **M**, **P** and **L**, then $\mathbf{X} = \{fast, Vfast, Mfast, LPfast, Lfast, Pfast, Lslow, slow, Pslow, Vslow, ...\}$ U 0, W, 1 is a term-domain of **X**.

It can be considered as an abstract algebra $\mathbf{AX} = (\mathbf{X}, \mathbf{C}, \mathbf{H}, \leq)$, where **H** is a set of linguistic hedges, which can be regarded as one-argument operations, $\leq$ is called a semantics-based ordering relation on **X** and **W**, 0, 1 is a set of constants in **X** with fast and slow being primary terms of **X** and **W**, 0, 1 being additional elements in **X** interpreted as the neutral, the least and the greatest ones, respectively. Denote by **hx** the result of applying an $h \in \mathbf{H}$ to $x \in \mathbf{X}$ and by $H(x)$ the set of all $u \in \mathbf{X}$ generated algebraically from x by using hedges in **H**, i.e. $H(x) = u$: $u = h_n ... h_1 x, h_1, ..., h_n \in \mathbf{H}$.

It is natural that there is a demand to transform fuzzy sets defined on a real interval $[a, b]$, which represents the meaning of terms in a term-domain **X**, into $[a, b]$ or, for normalization, into $[0, 1]$. This defines a mapping of the term-domain **X** into $[0, 1]$, called in the algebraic approach a semantically quantifying mapping. Now, we take these mappings in mind to define a notion of fuzziness measure. Let us consider a mapping f from **X** into $[0, 1]$, which preserves the ordering relation on **X**. Then, the "size" of the set $H(x)$, for $x \in \mathbf{X}$, can be measured by the diameter of $f(H(x)) \subseteq [0, 1]$. That is that this diameter will be considered as a fuzzy measure of the term x. Taking this model of fuzziness measure in mind, we may adopt the following definition:

Let $\mathbf{AX} = (\mathbf{X}, \mathbf{C}, \mathbf{H}, \leq)$ be a linear HA. An fm: $\mathbf{X} \rightarrow [0, 1]$ is said to be a fuzzy measure of terms in **X** if:

Definition 4. *For each $x \in \mathbf{X}$, the length of x is denoted by $|x|$, and defined as follows:*

1) if $x = c^+$ or $x = c^-$ then $|x| = 1$.
2) if $x = hx'$ then $|x| = 1 + |x'|$, for all $h \in \mathbf{H}$.

Proposition 1. *The fuzziness measure (fm) and the fuzziness measure of hedge h, denoted by $\mu(h)$, $\forall h \in \mathbf{H}$, with the following properties:*

1) $fm(hx) = \mu(h) \times fm(x)$ with $\forall x \in \mathbf{X}$;
2) $fm(c^+) + fm(c^-) = 1$;
3) $\sum_{-q \le i \le p, i \ne 0} fm(h_i c) = fm(c)$, $c \in \{c^+, c^-\}$;
4) $\sum_{-q \le i \le p, i \ne 0} fm(h_i x) = fm(x)$;
5) $\sum_{-q \le i \le -1} \mu(h_i) = \alpha$, $\sum_{1 \le j \le p} \mu(h_j) = \beta$, $\alpha + \beta = 1$, $(\alpha, \beta > 0)$

3 Algorithm

Table 1. The symbols used in the algorithm

Symbol	The meaning		Symbol	The meaning
D	transaction database		$D^{\sim}$	fuzzy transaction database
minsup	minimum support		minconf	minimum confidence
min_s	threshold language of minsup		min_c	threshold language of minconf
N	the total number of transaction data		M	the total number of items
d	the total number of managers		V	hedge "Very"
X	linguistics variable		M	hedge "More"
P	hedge "Possibly"		L	hedge "Less"
K	The number of fuzzy partitions in each items			

Input:

- D: The data set includes n quantitative transactions;
- Table qualitative: A set of m items with their importance evaluated by d managers;
- A pre-defined linguistic minimum support valuemin_s and linguistic minimum confidence valuemin_c.

Output: A set of fuzzy association rules.
Method: Includes 9 following general steps:
Step 1: Identify minsup, minconf from the pre-definedthreshold linguistic.

Transform min_sas **X** variables in HA. Including:

+ Calculate the fuzzy of variable **X**: $fm(\mathbf{X})$;
+ Identify fuzzy approximately of **X**: $I(\mathbf{X}) = [a, b]$;

+ The fuzzy average value of the variable $\mathbf{X}$:

$$gt(\mathbf{X}) = \frac{a+b}{2}; \tag{1}$$

+ Similarly, the linguistic confidence min_c;
+ Select linguistic thresholds (X, Y), respectively to the fuzzy value of (X, Y) as minsup, minconf: $minsup(\mathbf{X}) = gt(\mathbf{X})$; $minconf(\mathbf{X}) = gt(\mathbf{Y})$.

Step 2: Handling qualitative table: A set of m items with their importance evaluated by d managers

+ Calculate the fuzzy of linguistic variables;
+ Calculate the average offuzzy approximatelyqualitativeterms for all items.

$$kdt^{\sim}_{tb}(j) = \frac{1}{d} \times \sum_{i=1}^{d} (a(j)_i, b(j)_i)\ ;\ \text{(has the form: } [a_j, b_j]) \tag{2}$$

+ Calculate the average of fuzzy value for each item:

$$gtdt^{\sim}_{tb}(j) = \frac{a_j + b_j}{2}; \tag{3}$$

where: a_j and b_j are the values of $kdt^{\sim}_{tb}(j)$, which $kdt^{\sim}_{tb}(j) = [a_j, b_j]$

Step 3: Handling n quantitative transactions.

+ Transform the quantitative valueas A_j $(j = (\overline{1, m}))$ as $\mathbf{X}$ variables in HA $(\mathbf{X} \in \underline{\mathbf{X}})$, determined as follows:
  $\underline{\mathbf{X}}_{sl} = (X_{sl}, G_{sl}, H_{sl}, \leq)$, with: $G_{sl} = \{High, Low\}$, $(High = H, Low = L)$; $c^+ = \{H\}$; $c^- = \{L\}$; $H^+_{sl} = \{Very, More\}$; $H^-_{sl} = \{Less, Possibly\}$; (with $Very > More$; $Less > Possibly$)
  - Selection: Dom(sl); fm(H); fm(L); fm(V); fm(M); fm(L); fm(P);
  - Identify fuzzy approximately of X is $I(X)$, with $X \in \underline{\mathbf{X}}$
  - Transform the quantitative value of item into $[0, 1]$ respectively;
    With each $A_j \in [0, 1]$ that into fuzzy approximately $I(X)$, respectively;
+ Statistics of fuzzy partitions in $D^{\sim}$
+ Find the largest fuzzy partition as representative of each item j^{th}:

$$max\_count_j = \max(count_{ji}),\ \text{with } i = (\overline{1, K}); \tag{4}$$

Step 4: Calculate the fuzzy support of each item ($j = \overline{1, m}$), as:

$$sup(j) = \frac{gtdt^{\sim}_{tb}(j) \times max\_count_j}{N}; \tag{5}$$

where $gtdt^{\sim}_{tb}(j)$ is the qualitative value (calculated by formula (3), in step 2); max_count_j is the quantitative vaule (calculated by formula (4), in step 3); and N is the total number of transaction data, $N = |D|$.
Step 5: Filter out all items in $D^{\sim}$, such that: satisfied frequent item of minimum support: $sup(item) \geq minsup$.
Step 6: Establish Fuzzy FP-tree: establish Header table; establish FP-tree
Step 7: Calculate the fuzzy qualitative of n-itemset ($K \geq n \geq 2$).

+ Find out of all frequent itemsets (denote by n-itemset) from FP-tree;
+ Calculate the qualitative of n-itemset.

Step 8: Calculate the fuzzy support of each n-itemset.

+ Using the formula (5) - in step 4:

$$sup(n - itemset) = \frac{gtdt_{tb}^{\sim}(j) \times max\_count_j}{N};$$

+ Filter out all n-itemset, such that: satisfied frequent items of minimum support: $sup(n - itemset) \geq minsup$. $(n \geq 2)$

Step 9: Export rules, calculate the confidence and check with minconf. Using the following substeps:

+ Check the association rules from result of step 8, each n-itemset with items $(A_1, A_2, ..., A_n)$, $(n = \overline{2, M})$: A_1^... ^A_{i-1}^$A_{i+1}...A_n \rightarrow A_i$; $(i = \overline{1, M})$
+ Calculate the fuzzy confidence value of each possible fuzzy association rule as:

$$conf(A \rightarrow B) = \frac{sup(A \cup B)}{sup(A)}; \tag{6}$$

+ Select the satisfied fuzzy association rule of minimum confidence.
  During use of HA for fuzzy transaction database and quantify of linguitics, we view each element of HA is a fuzzy region. So, the process of creating fuzzy region based on the structure of HA will simple, intuitive, and more fficient.

4 An example

In this section, an example is given to illustrate the proposed algorithm.
Input: includes three data follows:

1. The data set includes six quantitative transactions, as show in Table 2.
2. The importance of the items is evaluated by three managers as shown in Table 3.
3. A pre-defined linguistic minimum support value min_s and linguistic minimum confidence value min_c.

Table 2. Data transactions (denoted by D)

TID	Items
1	(A, 3) (B, 4) (C, 2) (D, 3) (E, 7) (F, 2)
2	(A, 3) (B, 7) (D, 3) (E, 10) (F, 7)
3	(A, 2) (B, 10) (C, 5) (D, 2) (E, 10) (F, 5)
4	(B, 10) (C, 10) (E, 10) (F, 10)
5	(A, 7) (D, 7) (E, 7) (F, 10)
6	(A, 2) (B, 10) (D, 2) (E,10) (F,10)

Table 3. The item importance evaluated by three managers

Item	Manager 1	Manger 2	Manager 3
A	Important	Ordinary	Ordinary
B	Very Important	Important	Important
C	Ordinary	Important	Important
D	UnImportant	UnImportant	Very UnImportant
E	Important	Important	Important
F	Important	Important	Ordinary

Output: A set of fuzzy association rules.
Method: Includes 9 following general steps:
Step 1: Identify minsup, minconf from the pre-defined threshold linguistic

Identify parameters in HA: $\underline{\mathbf{X}} = (\mathbf{X}, \mathbf{G}, \mathbf{H}, \leq)$, with:
$\mathbf{G} = \{Low, High\}$; $c^+ = High$ (denoted by H); $c^- = Low$ (denoted by L); $\mathbf{H}^+ = \{Very, More\}$, $\mathbf{H} = \{Less, Possibly\}$; (with: $Very > More$; $Less > Possibly$) with: $fm(L) = 0.3$; $fm(H) = 0.7$; $fm(V) = fm(M) = fm(L) = fm(P) = 0.25$;

Identify fuzzy degree and fuzzy approximately of $\mathbf{X}$:

With the variable X contains $c^- =$ "Low":

+ $fm(VL) = 0.25 \times 0.3 = 0.075 \Rightarrow I(VL) = [0, 0.075] \Rightarrow I(VL)_{TB} = 3.75\%$
+ $fm(ML) = 0.25 \times 0.3 = 0.075 \Rightarrow I(ML) = [0.075, 0.15] \Rightarrow I(ML)_{TB} = 11.25\%$
+ $fm(PL) = 0.25 \times 0.3 = 0.075 \Rightarrow I(PL) = [0.15, 0.225] \Rightarrow I(PL)_{TB} = 18.75\%$
+ $fm(LL) = 0.25 \times 0.3 = 0.075 \Rightarrow I(LL) = [0.225, 0.3] \Rightarrow I(LL)_{TB} = \mathbf{26.25}\%$

Similar, with the variable X contains $c^+ =$ "$High$":

+ $fm(LH) = 0.25 \times 0.7 = 0.175 \Rightarrow I(LH) = [0.3, 0.475] \Rightarrow I(LH)_{TB} = 38.75\%$
+ $fm(PH) = 0.25 \times 0.7 = 0.175 \Rightarrow I(PH) = [0.475, 0.65] \Rightarrow I(PH)_{TB} = 56.25\%$
+ $fm(MH) = 0.25 \times 0.7 = 0.175 \Rightarrow I(MH) = [0.65, 0.825] \Rightarrow I(MH)_{TB} = \mathbf{73.75}\%$
+ $fm(VH) = 0.25 \times 0.7 = 0.175 \Rightarrow I(VH) = [0.825, 0.1] \Rightarrow I(VH)_{TB} = 91.25\%$

- Select minsupport with linguistic thresholds as "Less Low" (denoted by LL)

$$minsup = minsup(LL) = 26.25\%$$

- Select minconf with linguistic thresholds as "More High" (denoted by MH)

$$minconf = minconf(MH) = 73.75\%$$

Step 2: Handling qualitative table: A set of m items with their importance evaluated by 03 managers.

Identify parameters in HA: Denote:
I: Important; uI: UnImportant; O: Ordinary;
VI: Very Important; VuI: Very UnImportant;

$\underline{\mathbf{X}}_{qt} = (X_{qt}, G_{qt}, H_{qt}, \leq)$, with: $G_{qt} = \{Important, UnImportant\}$; $c^+ = Important$; $c^- = UnImportant$; $H^+_{qt} = \{Very, More\}$; $H^-_{qt} = \{Less, Possibly\}$; (with: $Very > More$; $Less > Possibly$).

Let: $W_{qt} = 0.5$; $fm(I) = 0.4$; $fm(uI) = 0.6$; $fm(V) = 0.3$; $fm(M) = 0.2$; $fm(L) = 0.3$; $fm(P) = 0.2$;

Should have: $fm(VI) = 0.3 \times 0.4 = 0.12 \Rightarrow I(VI) = [0.88, 1]$; $fm(VuI) = 0.3 \times 0.6 = 0.18 \Rightarrow I(VuI) = [0, 0.18]$; $fm(O) = 0.5 \Rightarrow I(O) = [0.25, 0.75]$;

Table 3 is converted into Table 4, where $kdt^{\sim}_{tb}$ is the average of fuzzy approximately qualitative; $gtdt^{\sim}_{tb}$ is the average of fuzzy value.

Table 4. The item importance evaluated by three managers

Item	**Manager 1**	**Manger 2**	**Manager 3**	$kdt^{\sim}_{tb}$	$gtdt^{\sim}_{tb}$
A	$[0.6, 1]$	$[0.25, 0.75]$	$[0.25, 0.75]$	$[0.367, 0.833]$	0.6
B	$[0.88, 1]$	$[0.6, 1]$	$[0.6, 1]$	$[0.693, 1]$	0.85
C	$[0.25, 0.75]$	$[0.6, 1]$	$[0.6, 1]$	$[0.483, 0.92]$	0.7
D	$[0, 0.6]$	$[0, 0.6]$	$[0, 0.18]$	$[0, 0.46]$	0.23
E	$[0.6, 1]$	$[0.6, 1]$	$[0.6, 1]$	$[0.6, 1]$	0.8
F	$[0.6, 1]$	$[0.6, 1]$	$[0.25, 0.75]$	$[0.483, 0.92]$	0.7

Step 3: Handling n quantitative transactions.

Transform the quantitative valueas A_j ($j = \overline{1, m}$) as X variables in HA ($X \in \underline{\mathbf{X}}$), determined as follows:

$\underline{\mathbf{X}}_{sl} = (X_{sl}, G_{sl}, H_{sl}, \leq)$, with:

$G_{sl} = \{c^-, c^+\}$, with: $c^+ = High$ (denoted by H); $c^- = Low$ (denoted by L); $H^+_{sl} = \{Very, More\}$; $H^-_{sl} = \{Less, Possibly\}$; (with: $Very > More$; $Less > Possibly$) ($Very$, $More$, $Less$, $Possibly$ denoted by: V, M, L, P respectively)

Let: $Dom(sl) = [0, 13]$; $fm(H) = 0.4$; $fm(L) = 0.6$; $fm(V) = 0.15$; $fm(M) = 0.25$; $fm(L) = 0.35$; $fm(P) = 0.25$;

Should have: $fm(VL) = 0.15 \times 0.6 = 0.09$; $fm(ML) = 0.25 \times 0.6 = 0.15$; $fm(PL) = 0.25 \times 0.6 = 0.15$; $fm(LL) = 0.35 \times 0.6 = 0.21$;

Because: $VL < ML < Low < PL < LL$, should: $I(VL) = [0, 0.09]$; $I(ML) = [0.09, 0.24]$; $I(PL) = [0.24, 0.39]$; $I(LL) = [0.39, 0.6]$;

Similar: $fm(VH) = 0.15 \times 0.4 = 0.06$; $fm(MH) = 0.25 \times 0.4 = 0.1$; $fm(PH) = 0.25 \times 0.4 = 0.1$; $fm(LH) = 0.35 \times 0.4 = 0.14$;

Because: $VH > MH > High > PL > LH$, should: $I(LH) = [0.6, 0.74]$; $I(PH) = [0.74, 0.84]$; $I(MH) = [0.84, 0.94]$; $I(VH) = [0.94, 1.0]$.

From: $Dom(sl) = 2, 3, 4, 5, 7, 8, 9, 10$ convert into [0, 1]

Converted into: $Dom(sl) = \{0.15, 0.23, 0.3, 0.38, 0.53, 0.61, 0.69, 0.76\}$

Because: $0.15, 0.23 \in [0.09, 0.24] \equiv ML$ should: 0.15 and $0.23 \in ML$;

Similar: 0.3 and $0.38 \in [0.24, 0.39] \equiv PL$; $0.53 \in [0.39, 0.6] \equiv LL$; 0.61 and $0.69 \in [0.6, 0.74] \equiv LH$; $0.76 \in [0.74, 0.84] \equiv PH$.

We tabulated transaction was fuzzy

Next, statistics of fuzzy partitions in $D^{\sim}$ (result from table 5)

Table 5. Database transaction was fuzzy (denoted by $D^{\sim}$)

TID	Fuzzy items
1	(0.23/A.ML) (0.3/B.PL) (0.15/C.ML) (0.23/D.ML) (0.53/E.LL) (0.15/F.ML)
2	(0.23/A.ML) (0.53/B.LL) (0.23/D.ML) (0.76/E.PH) (0.53/F.LL)
3	(0.15/A.ML) (0.76/B.PH) (0.38/C.PL) (0.15/D.ML) (0.76/E.PH) (0.53/F.LL)
4	(0.76/B.PH) (0.76/C.PH) (0.76/E.PH) (0.76/F.PH)
5	(0.53/A.LL) (0.53/D.LL) (0.53/E.LL) (0.76/F.PH)
6	(0.15A.ML) (0.76/B.PH) (0.15/D.ML) (0.76/E.PH)(0.76/F.PH)

Table 6. Statistics of fuzzy partitions

Fuzzy item	Count	Fuzzy item	Count
A.ML	0.76	D.ML	0.76
A.LL	0.53	D.LL	0.53
B.PL	0.30	E.LL	1.06
B.LL	0.53	E.PH	3.04
B.PH	2.28	E.LL	0.53
C.ML	0.15	F.ML	0.15
C.ML	0.38	F.LL	1.06
C.PH	0.76	F.PH	2.28

Using formula (4), find out the largest fuzzy partition (result from table 6) as representative of each item:

Table 7. Fuzzy item

Fuzzy item	Count	Fuzzy item	Count
E.PH	3.04	A.ML	0.76
F.PH	2.28	D.ML	0.76
B.PH	2.28	C.PH	0.76

Step 4: Calculate the fuzzy support of each item (1-itemset).

Using formula (5): For example with item $E.PH$:

+ fuzzy approximately of support: $([0.6, 1] \times 3.04)/6 = [0.304, 0.51]$;
+ fuzzy value of support: $(0.304 + 0.51)/2 = 0.41 = 41\%$.

Step 5: Filter out all items in $D^{\sim}$. Such that: satisfied frequent item of minimum support: $sup(item) \geq minsup$.

If: $sup(item) < minsup$ (with: $minsup = 26.25\%$, result at Step 1)

Then: remove item in table 8.

Step 6: Establish fuzzy FP-tree: see figure 1

Step 7: Calculate the fuzzy qualitative of n-itemset

Substep 7.1: Find out of all frequent itemsets (denote by n-itemset) from FP-tree (see Table 11)

Table 8. Fuzzy support

Item fuzzy	Fuzzy approximately	Fuzzy support
E.PH	(0.304, 0.51)	41%
F.PH	$(0.483, 0.92) \times 2.28/6$	27%
B.PH	$(0.693, 1) \times 2.28/6$	32%
A.ML	$(0.367, 0.833) \times 0.76/6$	7.6%

Establish Header table

Table 9. Header

Item fuzzy	Support
E.PH	41%
B.PH	32%
F.PH	27%

Table 10. Filter $D^{\sim}$

TID	Transaction
1	(0.76/E.PH)
2	(0.76/B.PH) (0.76/E.PH)
3	(0.76/B.PH) (0.76/E.PH) (0.76/F.PH)
4	(0.76/F.PH)
5	(0.76/B.PH) (0.76/E.PH) (0.76/F.PH)

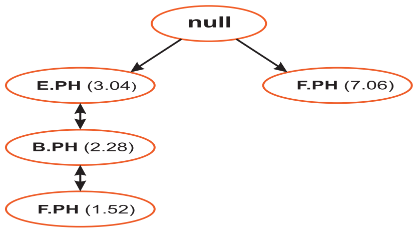

Fig. 1. Tree FP-tree

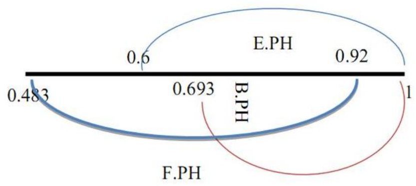

Fig. 2. Fuzzy approximately of 2 items on the qualitative attributes

Substep 7.2: Calculate the fuzzyqualitative of n-itemset.

For example with itemset BF: Fuzzyqualitative of B, F as [0.483, 0.92], [0.693,1], respectiely in figure 2.

Fuzzyqualitative of itemset BF as [0.693, 0.92]. Similar for other itemset.

Step 8: Calculate the fuzzy support of each n-itemset $(n \geq 2)$

Using the formula (5), example for itemset FE:

$$sup(F.PH, E.PH) = \frac{1.52 \times 0.76}{6} = 0.19 = 19\%$$

Thus, only itemset E.PH, B.PH satisfied frequent items of minsup.

Step 9: Export rules, calculate the confidence and check with minconf.

Result from step 8, we check two rules:

+ $E.PH \rightarrow B.PH$:

$$conf(E.PH \rightarrow B.PH) = \frac{sup(E.PH \cup B.PH)}{sup(E.PH)} = \frac{0.32}{0.41} = 78\%$$
$$> minconf(MH) = 73.75\%;$$

+ $B.PH \rightarrow E.PH$:

$$conf(B.PH \rightarrow E.PH) = \frac{sup(E.PH \cup B.PH)}{sup(B.PH)} = \frac{0.32}{0.323} = 99\%$$
$$> minconf(VH) = 91.25\%.$$

Table 11. Itemset should check the frequently

2-item	**3-item**
F.PH, B.PH: 1.52; F.PH, E.PH: 1.52; B.PH, E.PH: 2.28	F.PH, B.PH, E.PH: 1.52

Table 12. Fuzzy approximately of 2 items on the qualitative attributes

itemset	$kdt_{tb}^{\sim}$	$gtdt_{tb}^{\sim}$
F.PH, B.PH	(0.693, 0.92)	81%
F.PH, E.PH	(0.6, 0.92)	76%
B.PH, E.PH	(0.693, 1)	85%
F.PH, B.PH, E.PH	(0.693, 0.92)	81%

Table 13. Supportof itemset

Itemset	Support	Minsup = 26.25%
F.PH, E.PH	19%	unselected
F.PH, B.PH	21%	unselected
E.PH, B.PH	32%	selected
F.PH, B.PH, E.PH	21%	unselected

Result, we have 2 rules:

+ If E.PH then B.PH with a LL support and a MH confidence.
  If a Possibly High number of item E is bought, then a Possibly High number of item B is bought with a Less Low support and a MoreHigh confidence.
+ If B.PH then E.PH with a LL support and a VH confidence.
  If a Possibly High number of item B is bought, then a Possibly High number of item E is bought with a Less Low support and a Very High confidence.

5 Conclusion

The paper is an extension of the evaluation of fuzzy association rules was researched by Chien-Hua Wang and Chin-Pang Tzong [2], using algebras instead of fuzzy sets. The optimization of the parameters of quantitative semantic content in order to fit various problems will be discussed in our next papers.

References

1. Tran Thai Son and Nguyen Anh Tuan, Improve efficiency of fuzzy association rule using hedge algebra approach, Journal of Computer Science and Cybernetics, v.30, n.4, 397–408, 2014
2. Chien-Hua Wang and Chin-Tzong Pang, Finding Fuzzy Association Rules Using FWFP-Growth with linguistic Supports and Confidences, World Academy of Science, Engineering and Technology, 29, 1133-1141, 2009
3. Chien-Hua Wang, Chin-Tzong Pang and Sheng-Hsing Liu, Mining association rules uses fuzzy weighted FP-growth, Soft Computing and Intelligent Systems (SCIS) and 13th International Symposium on Advanced Intelligent Systems (ISIS), 2012 Joint 6th International Conference on, 13498461, 983 – 988, 2012
4. Tzung-Pei Hong , Chun-Wei Lin and Wen-Hsiang Lu, Lingguitic data mining with fuzzy FP-trees, Expert Systems with Applications 37, 4560-4567, 2010
5. Tzung-Pei Hong, Minh-Jer Chiang and Shyue-Liang Wang, Data Mining with Linguistic Thresholds, Int.Jcontemp. Math. Sciences, vol 7, n. 35, 1711 - 1725, 2012