Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 2, February 2014, pg.773 – 778

RESEARCH ARTICLE

Mining Multilevel Fuzzy Association Rule from Transaction Data

Urvi A. Chaudhary[1], Mahesh Panchal[2]

[123]Computer Department & Gujarat Technological University, Gujarat, India

[1] uapatel91@gmail.com; [2] mkhpanchal@gmail.com

Abstract:- Mining multilevel association rules in transaction dataset is most commonly and widely used in data mining. It is more challenging when some form of uncertainty like fuzziness is present in data or relationships in data. Present a model of mining multilevel association rules based on frequency. Due to this reason, the different minimum support at each level must be set a low value; otherwise, a lot of valuable patterns may not be found. We have employed fuzzy set concepts, multi-level taxonomy and different minimum supports to find fuzzy multilevel association rules in a given transaction data set. Apriori concept is used in model to find the item sets. The proposed model adopts a top-down progressively deepening approach to derive large itemsets. This approach incorporates fuzzy boundaries instead of sharp boundary intervals.

Keywords:- association rules, data mining, fuzziness, multilevel rules

I. INTRODUCTION

Apriori concept is used with fuzziness of data to find association rules. From hierarchical transaction dataset, find the strong association among intra-level & inter-level items.

A. *Overview*

Fuzzy logic was first developed by Zadeh [1] in the mid-1960s for representing uncertain and imprecise knowledge. It provides an approximate but Effective means of describing the behavior of systems that are too complex, ill-defined, or not easily analyzed mathematically. Fuzzy variables are processed using a system called a fuzzy logic controller. It involves fuzzification, fuzzy inference, and defuzzification. The fuzzification process converts a crisp input value to a fuzzy value. The fuzzy inference is responsible for drawing conclusions from the knowledge base. The defuzzification process converts the fuzzy control actions into a crisp control action. Fuzzy logic techniques have been successfully applied in a number of applications: computer vision, decision making, and system design including ANN training. The most extensive use of fuzzy logic is in the area of control, where examples include controllers for cement kilns, braking systems, elevators, washing machines, hot water heaters, air-conditioners, video cameras, rice cookers, and photocopiers.

Apriori is an influential algorithm in market basket analysis for mining frequent itemsets for Boolean association rules. The name of Apriori is based on the fact that the algorithm uses prior knowledge of frequent itemset properties. An Association rule mining is an important process in data mining, which determines the correlation between items belonging to a transaction database [3], [4]. In general, every association rule must satisfy two user

specified constraints: support and confidence. The support of a rule X □□□Y is defined as the fraction of transactions that contain XUY, where X and Y are disjoint sets of items from the given database [5], [6]. The confidence is defined as the ratio support (XUY)/support(X). Here the aim is to find all rules that satisfy user specified minimum support (□) and confidence values.

II. **LITERATURE REVIEW**

Fuzzy Logic was initiated in 1965, by Dr. Lotfi A. Zadeh[1] professor for computer science at the university of California in Berkley. Basically, Fuzzy Logic is a multi-valued logic that allows intermediate values to be defined between conventional evaluations like true/false, yes/no, high/low, etc. Fuzzy Logic starts with and builds on a set of user–supplied human language rules. Fuzzy Systems convert these rules to their mathematical equivalents. This simplifies the job of the system designer and the computer, and results in much more accurate representations of the way system behaves in real world. Fuzzy Logic provides a simple way to arrive at a definite conclusion based upon vague, ambiguous, imprecise, noisy, or missing input information.

A Fuzzy Set is any set that allows its members to have different grades of membership (membership function) in the interval [0,1]. In Fuzzy Logic, unlike standard conditional logic, the truth of any statement is a matter of degree. (e.g. How cold is it? How high shall we set the heat?) The degree to which any Fuzzy statement is true is denoted by a value between 0 and 1. Fuzzy Logic needs to be able to manipulate degrees of "may be" in addition to true and false.

A. *Mining Multilevel Association Rule using Fuzzy count Algorithm*[9].

Objective:

Find out intra level & inter level Fuzzy association rules.

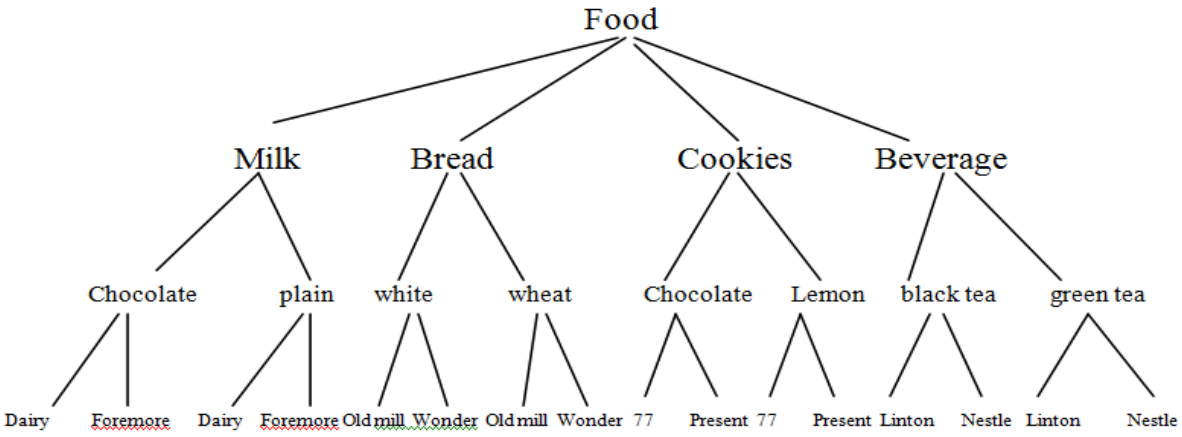

Fig 1. Transaction Dataset of Items

Trans_ID	List of items
T1	Wonder wheat bread, Limon black tea beverage, Old Mills white bread
T2	Limon black tea beverage, Nestle black tea beverage , Nestle green tea beverage, Old Mills white bread , Old Mills wheat brad
T3	Wonder wheat bread, Wonder white bread , 77 Lemon cookies, Nestle black tea beverage
T4	Present chocolate cookies, Dairy chocolate milk, Wonder white bread , Old Mills white bread , Limon black tea beverage, Nestle black tea beverage
T5	77 Lemon cookies, Wonder white bread , Nestle green tea beverage, Present chocolate cookies
T6	77 Lemon cookies, Wonder wheat bread, 77 Chocolate cookies, Limon black tea beverage
T7	Present chocolate cookies, Nestle black tea beverage , Old Mills wheat brad, Present chocolate cookies, 77 Lemon cookies
T8	Wonder white bread , Limon black tea beverage, Present lemon cookies, Old Mills wheat brad
T9	Present chocolate cookies, 77 Chocolate cookies, Linton green tea beverage, Wonder white bread , Old Mills white bread , 77 Lemon cookies
T10	77 Lemon cookies, Fore more plain milk
T11	Wonder wheat bread, Dairy chocolate milk
T12	Dairy chocolate milk

Table 1. Transaction Dataset

Firs each item is encoded as a sequence of digit. For example The item "foremore plain milk " in figure3.3 is encoded as '122'. In which the first digit '1' represents the code 'milk' at level 1, the second digit '2' represents the flavor 'plain (milk)' at level 2, and the third digit '2' represents the brand 'foremore' at level 3. Like that all items are encoded. Different minimum support setting should be used at different concept levels.

Algorithm steps:

1. Items of transaction dataset is encoded.
2. Define minimum threshold for each level.
3. Determine g ε {2, 3,.... }(Maximum item threshold).
4. Assume at level 1,
   minimum support = 5 %
   minimum confidence = 50%
5. Calculate sum of membership of item in transaction.
6. Compare fuzzy count with minimum threshold.
   find a set of single frequent items.
   A set of pair-wised frequent items.
   A set of strong association rules .
7. Repeat step-4, 5 & 6 at even lower levels until no frequent patterns can be found .

Trans_ID	List of items
T1	222, 411, 211
T2	411, 412, 422,211,221
T3	222, 212, 322, 412
T5	322, 212, 422, 311
T6	322, 222, 312, 411
T7	311, 412, 221, 311, 322
T8	212,411,321, 221
T10	322, 122
T11	222, 111
T12	111

Table 2. encoded transaction dataset

Outcomes:

After using above steps first find the frequent itemsets and then using these frequent itemsets find strong association rules.

B. *Mining Multilevel Association Rule using fuzzy sets Algorithm* [10]

Objective:

Find out intra level & inter level Fuzzy association rules.

Algorithm steps:

1. Items of transaction database is encoded.
2. Define same minimum threshold for each level.
3. Calculate sum of membership of fuzzy sets of item in transaction.
4. Compare fuzzy count with minimum threshold.
   find a set of single frequent items.
   A set of pair-wised frequent items.
   A set of strong association rules.
5. Repeat step-3 & 4 at even lower levels until no frequent patterns can be found .

In these algorithm first encoded items of transaction dataset same as the above method. Define same min_support threshold for each level. Then fuzzy count is find by using the membership function of product items.

Outcomes:

After using above steps first find the frequent itemsets and then using these frequent itemsets find strong association rules.

Motivation:

Mining multilevel fuzzy association rules at the same level and cross level.

Limitation:

- User define maximum item threshold, my improvement overcome the item threshold.
- User define maximum itemset threshold, my improvement overcome the item threshold.

III. PROPOSED APPROACH

Apriori concept is used with fuzziness of data to find association rules. From hierarchical transaction dataset, find the strong association among intra-level & inter-level items.

The Steps for proposed Algorithm are as follows:

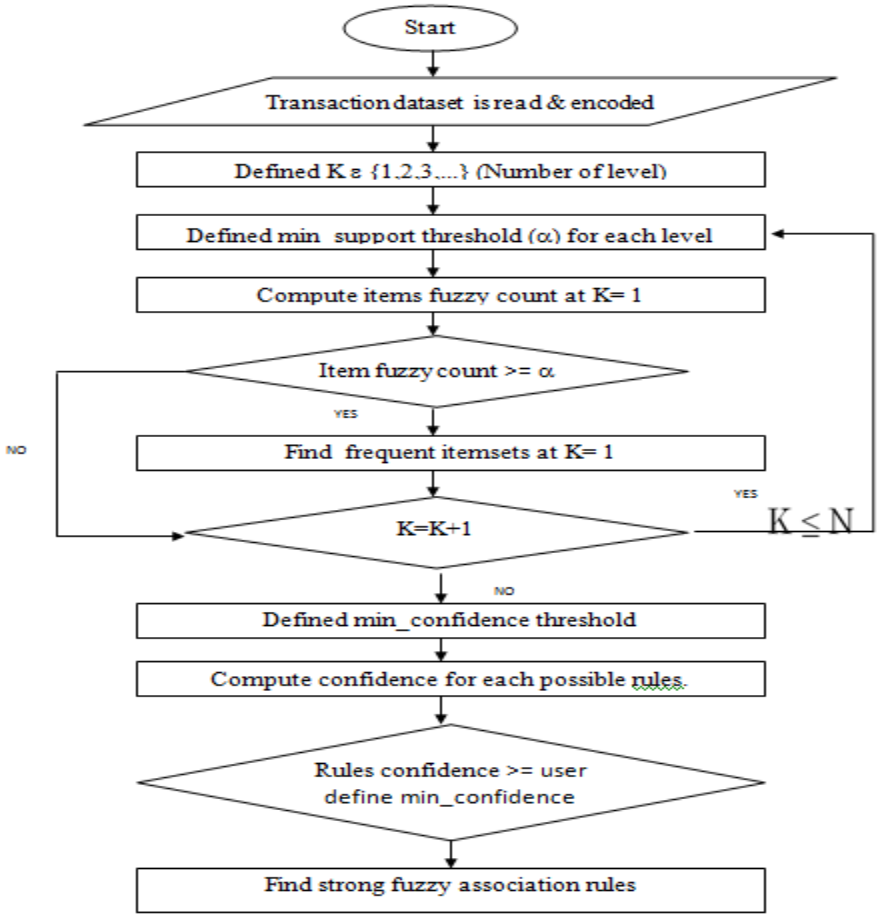

Figure 2: Mining Fuzzy Association rules Flow Chart

1. Take the transaction dataset as input.
2. Encode transaction dataset using a sequence of numbers and the symbol '*', with the ***I***th number representing the branch number of a certain item at levels.
3. Let k Ɛ{1,2,3,…}, where k is level number .
4. Determine minimum support threshold for each level (denoted by α) and minimum confidence threshold for association rules.
5. Compute items fuzzy count at k level.
6. Compare fuzzy count of items with min_support threshold.
   Item fuzzy count >= min_support.
7. Find frequent itemsets at k level.
8. This step is looking for possible candidate p-itemsets from Lp-1 at level k as there is no more candidate itemsets then increment the value of k.
9. Repeat the process from step-5.
10. No more frequent pattern can be found then go to step-11.
11. Calculate confidence for each possible association rules.
12. Compare confidence of rule with min_confidence.
    Association rule confidence >= min_confidence.
13. find strong fuzzy association rules.

IV. **CONCLUSIONS**

This paper, we have defined fuzzy set concepts, multiple level taxonomy and different minimum supports for each level and find fuzzy association rules in a given transaction data set. The algorithm works well with problems involving uncertainty in data relationships, which are represented by fuzzy set concepts. The proposed fuzzy mining algorithm can thus generate large itemsets level by level and then derive fuzzy association rules at same level, at different level and cross level from transaction dataset.

REFERENCES

1. L. A. Zadeh, " Fuzzy Sets", *Info. Control*,vol. 8, pp.338–353 ,1965.
2. LA Zadeh, " Fuzzy logic, neural networks and soft computing*", One-page ourse announcement* of CS 294-4, Spring 1993, University of California at Berkeley, 1992.
3. I. Ha",Y. Cai, and N. Cercone, **"**Data-driven of quantitative rules in relational databases," IEEE Tram. Knowledge and Data Eng**.,** vol. 5,
   pp.29-40, 1993.
4. R. Agrawal, T. Imielinski, and A. Swami, "Mining association rules between sets of items in large databases," In Proceeding ACM SIGMOD
   Conference, pp. 207-216, 1993.
5. J. Han, M. Kamber, "Data Mining: Concepts and Techniques," The Morgan Kaufmann Series, 2001.
6. R. Agrawal and R. Srikant. "Fast algorithms for mining association rules," In Proceedings of the 20th VLDB Conference, pp. 487-499, 1999.
7. K. R. Pardasani, Pratima Gautam "A Novel Approach For Discovery Multilevel Fuzzy Association Rules Mining, "Journal of Computing, Vol.2, Issue 3, March 2010.
8. Neural network ,Fuzzy logic, Genetic Algorithms, Synthesis & Applications : S. Rajasekaran & G.A .Vijayalakshmi Pai.
9. Pratima Gautam, Neelu Khare, & K. R. Pardasani, JOURNAL OF COMPUTING, VOLUME 2, ISSUE 1, JANUARY 2010, ISSN 2151-9617
10. Alireza Mirzaie Nejad Kousari and Ehsan Ghasemkhani World Applied Sciences Journal 17 (12): 1556-1565, 2012 ,ISSN 1818-4952

