2. Classification of Association Rule Mining Algorithms

Association rule mining algorithms can be divided in two basic classes; these are BFS like algorithms and DFS like algorithms [1]. In case of BFS, at first the minimum support is determined for all item sets in a specific level depth, but in DFS, it descends the structure recursively through several depth levels. Both of these can be divided further in two sub classes; these are counting and intersecting. Apriori algorithm comes under the counting subclass of BFS class algorithms. It was the first attempt to mine association rules from a large dataset. The algorithm can be used for both, finding frequent patterns and also deriving association rules from them. FP-Growth algorithm falls under the counting subclass of DFS class algorithms. These two algorithms are the popular example of the classical association rule mining.

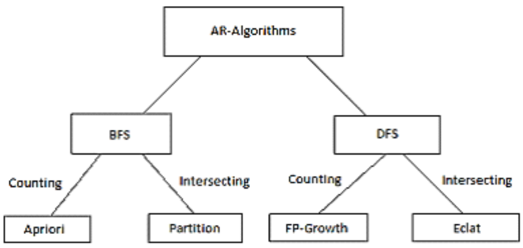

Figure 1: Classification of Mining Algorithm

3. Classical Association Rule Mining and Fuzzy Association Rule Mining

Classical association rule mining depends on the Boolean logic to transform numerical attributes into Boolean attributes by sharp partitioning of dataset. So that number of rules generated is low. It is inefficient in case of huge mining problems. In the classical association rule mining algorithms users have to specify the minimum support for the given dataset on which the association rule mining algorithm will be apply. But it is very much possible that the user sets a wrong minimum support value which can hamper the generation of association rules. And the setting of minimum support is not an easy task. If minimum support is set to a wrong value then there is a big possibility of combinatorial blow up of huge number of association rule within which many association rules will not be interesting. Fuzzy association rule mining first began in the form of knowledge discovery in Fuzzy expert systems. Instead of Boolean logic, a fuzzy expert system [2] uses a collection of fuzzy membership functions and rules [3]. The rules in a fuzzy expert system are usually of a form similar to the following:

"If it is raining then put up your umbrella"

Here if part is the antecedent part and then part is the consequent part [4]. This type of rules as a set helps in pointing towards any solution with in the solution set. But in case of Boolean logic every data attribute is measured only in terms of yes or no, in other words positive or negative. So it never allows us to have the diverse field of solutions. It always marginalizes the solutions; on the other hand fuzzy logic keeps broad ways of solutions open for the users. Many other fuzzy logic techniques are also use in fuzzy association rule mining [5]. Classical association rule mining uses the concept of crisp sets.

4. Crisp and Fuzzy Sets

Crisp Sets

Crisp sets have a well defined universe of members. A set can either be described by the list method (naming all the members) or by the rule method (properties that all members have to satisfy). Sets are denoted by capital letters, their members by lower-case letters. The list method is denoted as follows:

A={a1, a2,...,an}
For the rule method, we write:
B={b| b has propertiesP1, P2,...,Pn}

If the elements of a set are sets themselves, this set is referred to as a family of sets. {Ai| i∈I } Defines the family of sets where i and I are called the set identifier and the identification set. The family of sets is also called an indexed set. Any set A is called a subset of B if every member of A is also a member of B. This is written as A⊆B. If A⊆B and B⊆A, the two sets contain the same members and thus are equal. Equal sets are denoted by A=B, the contrary, namely unequal sets, are written as A≠B. If A⊆B and A≠B are true, this indicates that B contains at least one object that is not a member of A. In this case, A would be called a proper subset of B and written A⊂B. The empty set that contains no members is denoted by ∅.

Elements are assigned to the sets by giving them the values 0 or 1. Every element that shows a 1 is a member of the set. The number of elements that belong to a set is called its cardinality. All sets that have been created by the rule method might contain an infinite number of elements.

The set containing all members of set B that are not members of set A is called the relative complement of A with respect to set B, written B−A. If set B is the universal set, the compliment is absolute and denoted by the union of two sets A and B is a set containing all elements that are in A or in B, denoted by A∪B, whereas their intersection is a set that contains only those elements which are members of A and B, denoted by A∩B. Two elements are disjoint if they do not have any elements in common, that means, if A∩B=∅. A collection of disjoint subsets of A is called a partition on A if the union of those subsets makes the original set A. The partition is denoted by the symbol, formally ={Ai| i∈I, Ai⊆A}.

All of the operations union, intersection and complement apply to several rules.

At first, union and intersection are commutative, that means that the order of the operands does not affect the result:

A∪B=B∪A, A∩B=B∩A.

The second rule, called associativity, states that union and intersection can be applied pair wise in any order without changing the result:

A∪B∪C=(A∪B)∪C=A∪(B∪C),
A∩B∩C=(A∩B)∩C=A∩(B∩C).

Union and intersection both are idempotent operations because applying any of those operations on a set with itself will give the same set: A∪A=A, A∩A=A.

The distributive law is satisfied for both union and intersection in the following way: A∩(B∪C)=(A∩B)∪(A∩C), A∪(B∩C)=(A∪B)∩(A∪C).

DeMorgan's law constitutes that the complement of the union of two sets matches the union of their complements:

=∪, =∩. [6]

Fuzzy Sets

Fuzzy sets can generally be viewed as an extension of the classical crisp sets.

"Fuzzy sets are generalized sets which allow for a graded membership of their elements. Usually the real unit interval [0; 1] is chosen as the membership degree structure."[6]

Crisp sets are discriminating between members and nonmembers of a set by assigning 0 or 1 to each object of the universal set. By assigning values that fall in a specified range, typically 0 to 1, to the elements, fuzzy sets generalize this function. This evolved out of the attempt to build a mathematical model which can display the vague colloquial language. Fuzzy sets have proofed to be useful in many areas where the colloquial language is of influence. Let X be the universal set. The function $_A$ is the membership function which defines set A.

Formally: $_A$: X [0,1].

5. Fuzzy Association Rule Mining Algorithms

In the last few decades there has been a large number of research work already done in the field of fuzzy association rule mining. The concept of fuzzy association rule mining approach generated from the necessity to efficiently mine quantitative data frequently present in databases. Algorithms for mining quantitative association rules have already been proposed in classical association rule mining. Dividing an attribute of data into sets covering certain ranges of values, engages the sharp boundary problem. To overcome this problem fuzzy logic has been introduced in association rule mining. But fuzzy association rule mining also have some problems.. Classical association rule mining regarding the sharp partitioning. Following are some partitioning rule:

- Use of sharp ranges creates the problem of uncertainty. More precisely loss of information happens at the boundaries of these ranges. Even at the small changes in determining these intervals may create very unfamiliar results which could be also wrong.
- These partitions do not have proper semantics attached with them.

In fuzzy association rule mining the transformation of numerical attributes into fuzzy attributes is done using the fuzzy logic concept. In fuzzy logic attribute values are not represented by just 0 or 1. Here attribute values are represented within a range between 0 and 1[7]. According to this way, crisp binary attributes are converted to fuzzy attributes and by using fuzzy logic; we can easily resolve the above problems. The algorithms which are mostly use for fuzzy association rule mining are the fuzzy versions of Apriori algorithm. Apriori algorithm is slow and inefficient in case of large datasets. Fuzzy versions of Apriori algorithm would not be able to handle real-life huge datasets. Algorithms uses the principle of memory dependency like FP-Growth and its fuzzy versions are inadequate to deal with huge datasets. But these huge data sets can be easily managed by the partial memory dependent variant algorithms like ARMOR and. Ashish Mangalampalli, Vikram Pudi [12] proposed a new fuzzy association rule mining algorithm which will perform mining task on huge datasets efficiently and in fast. Their proposed algorithm has two-step processing of dataset. But before the actual algorithm there is preprocessing of dataset by fuzzy c-means clustering. Fuzzy partitions can be done on given data set so that every data point is a member of each and every cluster with a certain membership value. Main objective of the algorithm is to minimize the Equ $(1)^2$ (1)

Where m is any real number such that $1 \leq m < \infty$, μij is the degree of membership of xiin the cluster of j, xi is theith dimensional measured data,$cj$is the d-dimensional cluster center, and is any norm expressing the similarity between any measured data and the center. By this way corresponding fuzzy partitions of the dataset is generated where each value of numeric attributes are uniquely identified by their membership functions (μ). Depending upon the number of fuzzy partitions defined for an attribute, each and every existing crisp data is converted to multiple fuzzy data. This has the possibility of combinatorial explosion of generation of fuzzy records. So they have set a low threshold value for the membership function μ which is 0.1 to keep control over the generation of fuzzy records. During the fuzzy association rule mining process, the original data set is extended with attribute values within the range (0, 1) due to the large number of fuzzy partitions are being done on each of the quantitative attribute. To process this extended fuzzy dataset, some measures are needed which are based on t-norms [8], [9], [10]. In this way the fuzzy dataset E is created upon which the proposed algorithm will work. The dataset is logically divided into p disjoint horizontal partitions$P1$, 2,….,PP. Each partition is as large as it can fit in the available main memory. They have used the following notations,

- E=fuzzy dataset generated after pre-processing
- Set of partitions $P=\{P1,P2,\ldots.,PP\}$
- td[it] = tidlist of item set it
- μ = fuzzy membership of any itemset
- count[it] = cumulative μ of item set it over all partitions in which it has been processed
- d = number of partitions (for any particular item set it) that have been processed since the partition in which it was added

Byte-vector like data structure is used to represent fuzzy partitions of given data set. Each element of the byte-vector is nothing but the membership value (μ) of the item set. In a transaction the byte-vector cell which does not contain any item set, is assigned a value of 0. Initially the each byte-vector cell has the value of 0. Byte-vector representation of Tidlists is huge and could lead to incessant thrashing problem.

Quantitative association rules are defined over quantitative and categorical attributes [11]. The statement "70% of tertiary educated people between age 25 and 30 are unmarried" is one such example. In [11], the values of categorical attributes are mapped to a set of consecutive integers and the values of quantitative attributes are first discretized into intervals using equi-depth partitioning, if necessary, and then mapped to consecutive integers to preserve the order of the values/intervals. And as a result, both categorical and quantitative attributes can be handled in a uniform fashion as a set of <attribute, integer value> pairs. With the mappings defined in [11], a quantitative association rule is mapped to a set of boolean association rules. In other words, therefore, rather than having just one field for each attribute, there is a need to use as many fields as the number of different attribute values. For example, the value of a boolean field corresponding to <attribute1, value1> would be "1" if attribute1 has value1 in the original record and "0", otherwise [11]. After the mappings, the algorithms for mining Boolean association rules (e.g. [1-2]) is then applied to the transformed data set.

Let I = {i1, i2, …, i_m} be a set of binary attributes called items and T be a set of transactions. Each transaction t T is represented as a binary vector with t[k] = 1 if t contains item i_k and t[k] = 0, otherwise, for k = 1, 2, …, m. An association rule1 is defined as an implication of the form X Y where X I, Y I, and X Y =. The rule X Y holds in T with support defined as the percentage of records having both X and Y and confidence defined as the percentage of records having Y given that they also have X. For the mining algorithms to determine if an association is interesting, its support and confidence have to be greater than some user-supplied thresholds. A weakness of such approach is that many users do not have any idea what the thresholds should be. If it is set too high, a user may miss some useful rules but if it is set too low, the user may be overwhelmed by many irrelevant ones.

Furthermore, the intervals involved in quantitative association rules may not be concise and meaningful enough for human experts to obtain nontrivial knowledge. Fuzzy linguistic summaries introduced in [12-13] express knowledge in linguistic representation which is natural for people to comprehend. An example of linguistic summaries is the statement "about half of people in the database are middle aged." In contrast to association rules which involve implications between different attributes, the fuzzy linguistic summaries only provide summarization on different attributes. Although this technique can provide concise summaries which are nature for people to comprehend, there is no idea of implication in fuzzy linguistic summaries. As a result, this technique which provides a means for data analysis is not developed for the task of rule discovery. In addition to fuzzy linguistic summaries, the applicability of fuzzy modeling techniques to data mining has been discussed in [. Given a series of fuzzy sets, A1, A2, …, Ac, context sensitive Fuzzy C-Means (FCM) method is used to construct the rule-based models with the rules y is Ai if W1 and W2 and … and Wc where W1, W2, …, Wc are the regions in the input space that are centered around the "c" prototypes for i = 1, 2, …, c [13].

Nevertheless, the context-sensitive FCM method can only manipulate quantitative attributes and it is for this reason that this technique is inadequate to deal with most real-life databases which consist of both quantitative and categorical attributes.

6. Conclusion

Knowledge extraction in databases may be the method of extracting data within the type of interesting rules. These rules are domain specific. These rules reveal the association relationship among totally different data's that however a specific information items expounded to a different information item. So, we have a tendency to decision these rules as association rule. These rules are heuristic in nature. The method of extracting and managing these rules is understood as association rule mining. Association rule mining is a very important method in intelligent systems like Expert system. As a result of these intelligent systems solves domain specific issues. And this needs domain specific information. Association rule mining is essentially of two types. One is classical or crisp association rule mining and also the alternative one is fuzzy association rule mining. Classical association rule mining uses Boolean logic to convert numerical attributes into binary attributes by the assistance of sharp crisp partitions. However the utilization of sharp partitions creates the problem of uncertainty. Within which valuable information might become inconsistent over these sharp partitions. Another downside with classical association rule mining is, here a user need to offer a minimum support value for the mining purpose. And as we all know that we tend to humans are error prone. Any wrong setting of minimum support might find yourself in erroneous results. This will even cause the generation of large number of redundant rules further more as useless rules. So, it is very a difficult task of setting an accurate minimum support value manually. That is why classical association rule mining is time consuming and fewer accurate methods. Fuzzy association rule mining is comparatively a more recent idea. This uses the idea of fuzzy set theory for mining job. This survey paper represents a review of some of the existing fuzzy association rule mining methodologies.

References

[1] Hipp, Jochen; Guentzer, Ulrich; Nakhaeizadeh, Gholamreza: Algorithms for Association Rule Mining - A General Survey and Comparison. ACM SIGKDD Explorations Newsletter, Volume 2, Issue 1, 2000.
[2] Türks¸en, I.B. and Tian Y. 1993. Combination of rules and their consequences in fuzzy expert systems, Fuzzy Sets and Systems, No. 58,3-40, 1993
[3] http://www.cs.cmu.edu/Groups/AI/html/faqs/ai/fuzzy/part1/faq-doc-4.html
[4] L.A.Zadeh.Outline of a new approach to the analysis of complex systems and decision processes. IEEE Transactions on System, Man, and Cybernetics, Volume 3, Pages(s):28-44, January, 1973.
[5] Delgado, Miguel: Fuzzy Association Rules: an Overview. BISC Conference, 2003.
[6] Bakk. Lukas Helm, Dr. Michael Hahsler: Fuzzy Association Rules An Implementation in R, Master thesis, Vienna University of Economics and Business Administration.
[7] Ashish Mangalampalli, Vikram Pudi: Fuzzy Association Rule Mining Algorithm for Fast and Efficient Performance on Very Large Datasets. FUZZ-IEEE 2009, Korea, ISSN: 1098-7584, E-ISBN: 978-1-4244-3597-5, Page(s): 1163 – 1168, August 20-24, 2009.
[8] De Cock, M., Cornelis, C., Kerre, E.E.: Fuzzy Association Rules: A Two-Sided Approach. In: FIP, Page(s): 385-390, 2003.
[9] Yan, P., Chen, G., Cornelis, C., De Cock, M., Kerre, E.E.: Mining Positive and Negative Fuzzy Association Rules. In: KES, Springer, Page(s): 270-276, 2004.
[10] Verlinde, H., De Cock, M., Boute, R.: Fuzzy Versus Quantitative Association Rules: A Fair Data-Driven Comparison. IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics, Volume 36, Page(s): 679-683, 2006.
[11] R. Srikant, and R. Agrawal, "Mining Quantitative Association Rules in Large Relational Tables," in Proc. of the ACM SIGMOD Int'l Conf. on Management of Data, Monreal, Canada, June 1996, pp. 1-12.
[12] R.R. Yager, "On Linguistic Summaries of Data," in, pp. 347-363.
[13] R.R. Yager, "Fuzzy Summaries in Database Mining," in Proc. of the 11th Conf. on Artificial Intelligence for Application, Los Angeles, California, Feb. 1995, pp. 265-269