Volume 4, Issue 7, July 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

# A Review on the Present State-of-The-Art of Association Rule Mining

**V. Aruna[1], Dr. V. Janaki[2], Dr. Avula Damodaram[3]**
[1]Associate Professor, Aurora's Scientific Technological and Research academy, Banglaguda Hyderabad, India
[2]Professor Department of Information Technology in Vaagdevi Engineering College, Warangal, India
[3]Director, University Academic Audit Cell JNTU Kukatpally, Hyderabad, India

***Abstract-- Most of the data mining applications used by real world enterprises have mechanisms for frequent item set mining (FIM). The quality FIM is employed using statistical measures such as support and confidence. Association Rule Mining (ARM) is used to generate association rules from frequent item sets. Such rules encapsulate the latent relationship among objects in the real world. The bottom line of this indicates that association rule mining has very high utility in garnering business intelligence and make highly informed decisions. The ARM is of many types and it can be done on various kinds of data as well. This paper reviews the present state-of-the-art of association rule mining and algorithms. It also throws light into ARM with spatial data, stream data besides MapReduce programming for leveraging parallel processing. Boolean and maximal Boolean association rule mining, frequent item set mining, maximal frequent item set mining are explored. Various algorithms for association rule mining with underlying data on which they operate are discussed.***

***Keywords-- Data mining, frequent item set mining, maximal frequent item set mining, Boolean association rule mining***

## I. INTRODUCTION

Data mining is the process of discovering trends which were latent in the data sources. Frequent item set mining has been around for many years. It is an essential mining activity. Frequent item sets can provide user interesting facts that form business intelligence which helps in making expert decisions. Hidden associations from data of any domain can become actionable knowledge. Association rules can be mined from frequent item sets. The quality of frequent item sets can be measured using statistical measures like support and confidence. Let D be a transactional database. Let I be the set of items present in database D. X is a frequent item set present in I. The support is the measure which is computed as number of transactions with A and B (assuming A and B are items in I) divided by total number of records. This will help in filtering the frequent item sets for bringing about quality. In the same fashion, the statistical measure confidence is computed as number of transactions that contain A and B divided by the total number of records with A. These two measures can be used together to generate frequent item sets that are used to produce high quality association rules.

Association rules express the relation between objects. They are used to understand the relationship among objects. There are algorithms used to generate association rules. The process of generating them is called association rule mining (ARM). With ARM it is possible to analyze customer behavior and take intelligent decisions. Based on the values handled the association rules are of two types namely Boolean association rules and quantitative association rules. Based on the level of abstraction involved, they are divided into single level association rules and multilevel association rules. Based on the dimensions they are categorized into single-dimensional association rules and multidimensional association rules. The general form of association rule is Body => Head [Support, Confidence]. For instance A=>B [0.5%, 60%] is a rule where A and B are associated items with 50% support and 60% confidence. The support measure indicates the frequency of item set while the confidence denotes the strength of it. Generally minimum support and maximum confidence are provided by domain experts so that the generated association rules will have highly interested business intelligence. If the support threshold is low, it can produce many valid rules. If that is high, it produces few frequent item sets. If the confidence threshold is high, it produces few rules but with high accuracy. If the confidence threshold is low many rules are produced with less accuracy. Typical values for best results are support between 2 and 10% while the confidence between 70-90%.

Association rule mining is the two step process in which first step focuses on mining frequent item sets while the second steps generates association rules based on the threshold of the measures such as support and confidence. Out of the two steps the first step differs from algorithm to algorithm while the second step is same for all and it is simple too. Figure 1 shows the algorithms that are used in the real world for mining association rules on both quantitative and Boolean data. They are also different for certain and uncertain data.

 

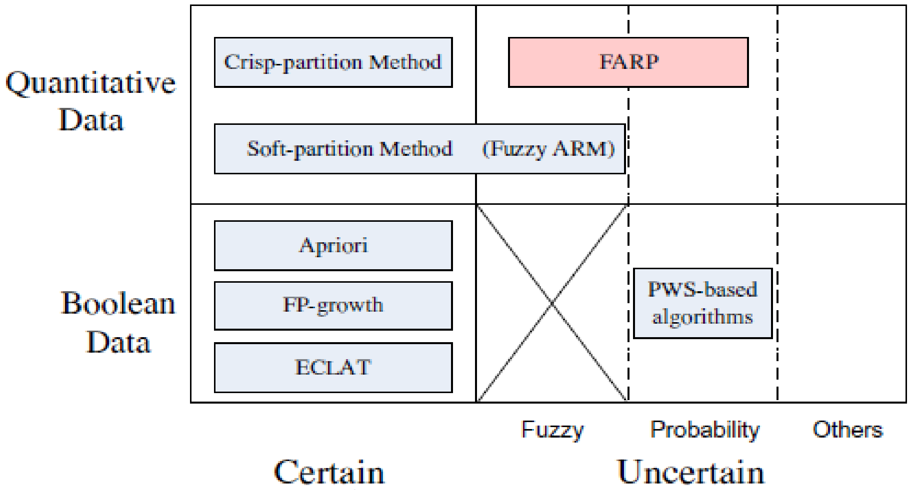


Fig.1. Categories of association rule mining algorithms [1]

As can be seen in Figure 1, it is evident that Apriori, FP-Growth and ECLAT are the algorithms that mine association rules on Boolean data which is certain. Crisp partition method and fuzzy ARM are the algorithms for mining rules on quantitative and certain data. FARP proposed in [1] works on quantitative uncertain data while the PWS-based algorithms work on uncertain Boolean data.

Our main contribution in this paper is to review the present state-of-the-art of association rule mining and provide research insights with respect to various kinds of mining association rules and the underlying methods or algorithms. The remainder of the paper is structured as follows. Section 2 focuses on frequent item set mining. Section 3 throws light into frequent closed item set mining. Section 4 focuses on the maximal frequent item set mining. Section 5 provides information on frequent item set mining over stream data. Section 6 focuses on distributed data processing for mining frequent item sets. Section 7 and 8 are on Boolean and maximal Boolean association rule mining respectively while section 9 concludes the paper.

## II. FREQUENT ITEMSET MINING

Having understood various kinds of item set mining in the previous section, this section focuses on various approaches used in the literature for FIM. Grahne et al. [2] introduces a new technique known as FP-array to improve the performance of FP-tree based algorithms. The technique works well with all sparse data sets. Moreover it can be combined with many optimization techniques. The technique helps in mining both frequent item sets and also maximal frequent item sets. Baralis et al. [3] studied FIM on relational databases. They introduced index support for FIM. Their approach solved the problem of mining frequent item sets effectively on RDBMS. I-tree index is the structure proposed by Baralis et al. which is efficient in terms of response time and memory usage. Wang and Wu [4] studied the problem of approximate inverse frequent item set mining. The main focus of the work was on approximation, privacy and complexity. The approximation algorithms are used to generate synthetic transactional databases and approximation of privacy leakage confidence level.

Bashir et al. [5] proposed many implementation techniques and strategies for best FIM. The efficient implementations of AFOPT, FP-Growth and Patricia Mine on various benchmark datasets of UCI machine learning repository. Later Liu et al. [6] introduced frequent item set mining that exhibits two level counting and heuristic in nature. It is named HTLC which improves the technique used for item set generation. Both traversal and support counting approaches are improved. Thus the HTLC performs better than many other Apriori-like algorithms.

Luo and Zhang [7] proposed a hybrid algorithm named “DiffsetHybrid” for efficient mining of frequent item sets. The algorithm is based on vertical database layout. It uses the notion of diffset that is the difference between two tidsets. For sparse datasets the hybrid algorithm has outperformed the algorithm based on tidset. Khare et al. [8] studied multi-dimensional fuzzy association rule mining. They proposed an algorithm that works on database with fuzzy transactions besides employing aproiri property for pruning. The algorithm brings about latent relationships among the dimensions in a relational database.

Frequent item sets generated as part of data mining could be sensitive in nature. Protecting the privacy There Gkoulalas-Divanis et al. [9] proposed a border-based approach for hiding sensitive frequent item sets. This is achieved by extending original database for optimal results. Generating item sets that not only reflect user interest, but also user-friendly is an important activity. Lim [10] introduce a new approach known as visual frequent item set mining (VFIM) which is an alternative to apriori-like approach. The VFIM approach enables domain experts to involve in generating frequent item sets interactively. The formal interaction between the program and the human expert makes this approach flexible and robust.

 

Gao and Wu [11] focused on probabilistic or uncertain data for frequent item set mining. They proposed an algorithm named "P-Apriori" for generating top-k probabilistic frequent item sets. The results will be rendered in the descending order of confidences. Polynomial algorithm is employed for testing candidates generated. Based on the parameters used such as k value, minimum support and minimum confidence, the algorithm produces results. Stanišić and Tomović [12] proposed a cross validation model for frequent item set mining. The cross validation procedure is coupled with Apriori algorithm for improving the quality of frequent item set generation. The execution time of proposed algorithm is reduced when compared with traditional Apriori. Stanisic and Tomovic [13] presented a mathematical solution for generating frequent item sets by using linear algebra method. The algorithm is based on apriori property.

### 2.1 Constraint Item Set Mining

Constraints can be applied to frequent item set mining to improve the quality of results. Leung et al. [14] studied constrained frequent item set mining on uncertain data. The uncertain data might be in the form of streaming data. Leung et al. proposed tree-based algorithms to mine frequent item sets efficiently under the constraints provided by user. Omari [15] proposed a temporal measure to obtain more meaningful frequent item sets. This measure can be used to improve search strategy employed by apriori in order to search from recent transactions. The measure helps in extracting more useful and user-interested patterns.

### 2.2 Deep Web vs. Frequent Item Set Mining

With the advent of Web 2.0, new features were available to built interactive web applications. This enabled a new way of information dissemination known as deep web which has limited interface for querying. Liu and Agrawal [16] proposed an approach that helps in extracting frequent item sets from deep web. This approach has proved 95% accuracy in mining frequent item sets.

### 2.3 FIM for Analyzing Binary Datasets

Binary datasets can also be subjected to frequent item set mining. Bhadoria et al. [17] proposed a pattern growth method based on hyper structure. The proposed model works for both binary datasets and also uncertain datasets. With respect to uncertain data, it employs clique based candidate pruning. This is achieved by extending H-mine algorithm. In case of sparse datasets H-mine performs well while GP-Growth performance is better in other cases. On all the tested datasets, H-Mine achieves better performance over Apriori algorithm. For large number of transactions and data items the H-mine algorithm is scalable.

### 2.4 FIM on Spatial Data

Spatial database holds the details of real world objects such as cities, roads, buildings, rivers etc. This database is very complex and mining frequent item sets (frequently co-located items) is a challenging task. Cheng et al. [18] proposed a new method for spatial mining of co-located frequent objects. The algorithm named Maximal Clique Enumeration Based on Ordered Star Neighborhood (MCEBOSON) to make data mining possible by transactionalization of spatial Boolean data. MCEBOSON algorithm combined with Apriori and FP Tree could not perform well when compared with MCEBOSON alone.

## III. FREQUENT CLOSED ITEM SET MINING

Frequent closed item sets are the item sets that have no immediate supersets with same support as the item sets in question. Closed frequent item sets uniquely reflect frequency of all item sets. Zaki and Hsiao [19] proposed an algorithm named "CHARM" for frequent closed item set mining. It makes use of dual itemset-tidset search tree coupled with a hybrid search. Thus this algorithm is capable of skipping many levels. Moreover it also reduces memory usage using the notion of diffsets. It also removes "nonclosed" sets using fast-hash based approach. For generation and visualization of rules CHARM-L algorithm was proposed by Zaki and Hsiao which makes the results intuitive.

Wang et al. [20] proposed an algorithm named TFP for extracting top-k frequent closed item sets. The specialty of this algorithm is to present top-k frequent closed item sets without using minimum support measure. Wang et al. also introduced the notion of descendent_sum and closed node count for pruning FP-tree effectively during and after construction of the tree. Moreover two novel strategies named "bottom-up" and "top-down" were introduced to speed up the mining process. This algorithm outperforms other efficient algorithms such as CHARM and CLOSET+ [21].

Closed frequent item sets provide many advantages but they can be made better using constraints. Wang et al. [22] incorporated constraints into the generation of frequent closed item sets. In fact they improved CLOSET+ and named it as LCLOSET. The LCLOSET is coupled with gradient mining algorithm named "FCCGM" for effectiveness. Later on Lucchese et al. [23] presented a scalable algorithm for closed FIM. The algorithm is fast and memory efficient. It make use of bitwise vertical representation of database besides using divide-and-conquer approach. It has several optimizations at various levels so as to make it more effective. This algorithm outperforms CLOSET and FP-CLOSE algorithms.

Mining closed frequent item sets from large databases and high dimensional database is challenging. Fu et al. [24] proposed an algorithm to address the challenges. Concept lattice structure is used to generate frequent item sets. Thus interpretation of mining results can be improved as closed frequent item sets reveal significant knowledge. The algorithm also uses cluster analysis in order to have informative item sets. Li et al. [25] an efficient algorithm that can generate closed frequent item sets. The algorithm makes use of FP-tree that scans database only once. For every node in the FP-tree a link is constructed and the pruning of the link is used while generating frequent closed item sets.

## IV. MAXIMAL FREQUENT ITEM SET MINING

When an item set X is frequent with no frequent superset, such X is said to be maximally frequent. Apriori [26] is the base line algorithm for generating frequent item sets. It employs breadth-first approach to discover frequent item sets at level L1 before going to L2. It also obtains support by generating and counting nodes explicitly. Later MaxMiner [27] came into existence which also used breadth-first traversal. The traversal also includes search space. However, it could perform lookaheads in order to prune unwanted branches so as to improve performance. MaxMiner uses this approach to reduce the number of passes over transactional database. DepthProject [28] is another maximal pattern method which uses combination of depth-first traversal and superset pruning besides dynamic reordering of child nodes. This will reduce search space and improves performance. MaxEclat and MaxClique [29] are the other maximal pattern methods available. They break subset lattice of a tree into small pieces and use a bottom-up Apriori fashion along with vertical data representation. These two algorithms limit future mining flexibility as they depend on pre-processing. Pincer-Search [30] also has pre-processing step. The VIPER algorithm proposed in [31] uses vertical layout and outperforms other algorithms that use horizontal layout. However, this algorithm is not suitable when patterns are too long as it returns the entire set of frequent items. Other vertical methods for mining maximal frequent item sets are explored in [32], [33] and [34]. Burdick et al. [35] developed a novel algorithm known as MAFIA which makes use of item set lattice for depth-first traversal. The traversal of the algorithm employs good pruning mechanisms. The search strategy is coupled with database represented as vertical bitmap which has provision for compression. When compared with DepthProject algorithm MAFIA is much better in terms of execution time.

## V. FREQUENT ITEM SET MINING OVER STREAM DATA

Frequent item set mining on streaming data which comes live is relatively new phenomenon. Qu et al. [47] explored possibilities of FIM over stream data. The stream data has specific characteristics such as fast arrival rate, covers wide territory, and time-varying. In task based approach general FIM, maximal FIM, and closed FIM on the streaming data was focused. With respect to data based algorithm, sampling, histogram, wavelet transforming, hashing, sketching and load shedding concepts were discussed with respect to FIM on streaming data. Memar and Deypir [36] proposed a block based approach for data streams to mine frequent item sets. The notion of sliding window is used along with blocked bit sequence technique in order to improve the mining time and sliding process.

Hou et al. [37] proposed an algorithm named FI-PS to generate frequent item sets on streaming data in the presence of concept drifts. As concept drifts, in case of streaming data, cause problems to mining, the proposed algorithm has provision for withstanding concept drifts. It is achieved as the algorithm is adaptive in terms of pre-computation of negative boundaries and sampling period. Thus the algorithm FI-PS outperforms other algorithms such as FDPM and LossyCounting.

## VI. DISTRIBUTED DATA PROCESSING FOR FREQUENT ITEM SET MINING

Frequent item set mining becomes difficult when there is huge amount of data to be processed. As the databases are growing day by day it is essential to speed up the processing. One approach to do this is distributed data processing or parallel processing with multiple processors. Xiaoyun et al. [38] proposed an algorithm for mining frequent item sets through parallel processing. The algorithm is named “HPFP-Miner” which is based on FP-Growth. Its communication overhead is little as it can efficiently partition. The scalability and performance of HPFP-Miner is higher than that of FP-Growth and F-Miner.

Chen et al. [39] proposed an algorithm for closed frequent item set mining based on MapReduce programming which is a new model for distributed programming environment. It uses certain steps like parallel counting, construction of frequent list and grouping, and parallel mining of patterns. The algorithm uses greedy approach in order to balance computation burdens. Another MapReduce solution for generating frequent item sets was proposed by Kov´acs and Ill´es [40]. Their algorithm is named as Porting Apriori algorithm tailored for MapReduce programming. The solution provides a reducer component that is cost effective and reduces the network traffic.

## VII. MINING BOOLEAN ASSOCAITION RULES

Based on the Boolean association rules, Zi-Yang and Guo-Hua [41] proposed privacy-preserving quantitative association methods (defined in introduction). These methods perform association rule mining with privacy of data preserved. Ke et al. [42] proposed information – theoretic approach for quantitative association rule mining. It is something different from Boolean association rule mining. The Boolean ARM is based on Boolean attributes while the quantitative ARM is based on attributes that hold quantitative values. To overcome the problems associated with huge number of item sets in case of quantitative attributes, Ke et al. used information-theoretic approach which avoids unnecessary or useless combinations. Prakash et al. [43] used multiple criteria in order to achieve Boolean association rule mining. It helps in counting support and confidence of suppliers based on given criteria. The proposed supplier selection criteria with Boolean association rule mining is useful in selecting eligible suppliers.

 

Angiulli et al. [44] explored the problem of finding complexity in association rule mining. The research was both on quantitative and Boolean association rules. Besides, they also used measures like laplace, gain, confidence and support. Intan [45] proposed an algorithm for association rule mining based on fuzzy sets. Market basket analysis is made using the proposed algorithm which is more intuitive when compared with Apriori algorithm. An interesting fact pertaining to Boolean association rule mining is that Apriori algorithm is used for mining frequent item sets in order to generate Boolean association rules. Pei et al. [1] focused on mining Boolean association rules from probabilistic quantitative databases. They proposed support and confidence measures specifically for such databases in order to quantify association rules. They proposed an algorithm named FARP to discover fuzzy association rules based on the statistical measures introduced.

## VIII. MINING MAXIMAL BOOLEAN ASSOCIATION RULES

Maximal Boolean association rules can improves the quality of the rules. Towards this end Jiang et al. [46] studied the property such as maximal Boolean association rules besides proposing an algorithm that can extract frequent item sets. From the experiments it was known that opened and closed frequent item sets can be used to mine maximal Boolean association rules. The rules thus generated are more meaningful. The advantage of maximal Boolean association rules is that they reduce decrease number of association rules without losing significant association rules. This will help in getting quality patterns that help in well informed decision making.

## IX. CONCLUSION AND FUTURE WORK

In this paper, we study association rule mining and frequent item set mining with respect to various kinds of data and algorithms used there in. There are many algorithms found in the literature that act on Boolean and quantitative data with certain and uncertain kinds. These algorithms are meant for producing frequent item sets that can be used to generate association rules. The paper also throws light into constraint item set mining, item set mining for deep web, frequent closed item set mining, maximal frequent item set mining, frequent item set mining for stream data, frequent item set mining with MapReduce kind of distributed data processing framework. It also throws light into Boolean and maximal Boolean association rule mining. Our future work includes building mining techniques for discovering frequent closed item sets to generate maximal Boolean association rules for efficient decision making.

## REFERENCES

[1] Bin Pei, Suyun Zhao, Hong Chen, Xuan Zhou and Dingjie Chen, "FARP: Mining fuzzy association rules from a probabilistic quantitative database", Information Sciences, Elsevier, 2013, p242-260.

[2] Go¨ sta Grahne and Jianfei Zhu. (2005). Fast Algorithms for Frequent Itemset Mining Using FP-Trees. *ieee*. p1347-1362.

[3] lena Baralis,Tania Cerquitelli and Silvia Chiusano. (2005). Index Support for Frequent Itemset Mining in a Relational DBMS. *ieee*.

[4] Yongge Wang, Xintao Wu and UNC Charlotte. (2005). Approximate Inverse Frequent Itemset Mining: Privacy, Complexity, and Approximation∗. *ieee*.

[5] Shariq Bashir, Muhammad Shuaib, Yasir Sultan and Dr. A. Rauf Baig. (2006). Improving Frequent Itemset Mining Algorithms Performance Using Efficient Implementation Techniques: A Benchmark Solution. *ieee*. p257-262.

[6] Feng Liu,FengZhan Tian and QiLiang Zhu. (2006). Frequent Itemset Mining Based on Heuristic Two Level Counting. *ieee*.

[7] KE LUO, XUE-MAO ZHANG. (2007). AN EFFICIENT FREQUENT ITEMSET MINING ALGORITHM. *ieee*. p1-6.

[8] Neelu Khare,Neeru Adlakha and K. R. Pardasani. (2009). An Algorithm for Mining Multidimensional Fuzzy Assoiation Rules. *ieee*. p1-5.

[9] Aris Gkoulalas-Divanis,Vassilios S. Verykios,. (2009). Exact Knowledge Hiding through Database Extension. *ieee*. p1-15.

[10] SeungJin Lim. (2009). On A Visual Frequent Itemset Mining. *ieee*. p1-6.

[11] Feng Gao, Chengrong Wu. (2011). Mining Frequent Itemset from Uncertain Data. *ieee*. p1-5.

[12] Predrag Stanišić,Savo Tomović. (2012). Frequent Itemset Mining Using Two-Fold Cross- Validation Model. *ieee*. p1-4.

[13] Predrag Stanišić,Savo Tomović (2013). Frequent Itemset Mining as Set Intersection Problem. *ieee*. p1-4.

[14] Carson Kai-Sang Leung,Boyu Hao, and Fan Jiang. (2010). Constrained Frequent Itemset Mining from Uncertain Data Streams. *ieee*. p1-8.

[15] Tantan Liu , Gagan Agrawal. (2011). Active Learning Based Frequent Itemset Mining Over the Deep Web. *ieee*. p1-12.

[16] Mina Memar,Mahmood Deypir,Mohammad Hadi Sadreddini,Seyyed Mostafa Fakhrahmad. (2011). A Block-Based Approach for Frequent Itemset Mining over Data Streams. *ieee*. p1-5.

[17] Robin Singh Bhadoria,Ram Kumar and Manish Dixit. (2011). Analysis on Probabilistic and Binary Datasets through Frequent Itemset Mining. *ieee*. p1-5.

[18] Yang Cheng, Zhang Tianjun, Lu Junli. (2013). A Maximal Clique Enumeration Based on Ordered Star Neighborhood for Co-location Patterns. *ieee*. p1-4.

 

[19] Mohammed J. Zaki, Member, IEEE, and Ching-Jui Hsiao. (2005). Efficient Algorithms for Mining Closed Itemsets and Their Lattice Structure. *ieee*. p462-478.
[20] Jianyong Wang, Jiawei Han,Ying Lu, and Petre Tzvetkov. (2005). TFP: An Efficient Algorithm for Mining Top-K Frequent Closed Itemsets. *ieee*. p652-663.
[21] J. Pei, J. Han, and R. Mao, “CLOSET: An Efficient Algorithm for Mining Frequent Closed Itemsets,” Proc. 2000 ACM-SIGMOD Int’l Workshop Data Mining and Knowledge Discovery (DMKD ’00), pp. 11- 20, May 2000.
[22] Jianyong Wang,Jianyong Wang and Jian Pei. (2006). Closed Constrained Gradient Mining in Retail Databases. *ieee*. p764-769.
[23] laudio Lucchese, Salvatore Orlando, and Raffaele Perego. (2006). Fast and Memory Efficient Mining of Frequent Closed Itemsets. *ieee*. p21-35.
[24] Huaiguo Fu, M´ıche´al ´O Foghl´u and Willie Donnelly. (2007). Frequent closed informative itemset mining. *ieee*. p1-5.
[25] Shengwei Li, Lingsheng Li, Chong Han. (2003). Mining Closed Frequent Itemset based on FP-Tree. *ieee*. p1-4.
[26] R. Agrawal, R. Srikant: Fast Algorithms for Mining Association Rules. ***Proc. of the 20th Int? Conference on Very Large Databases,*** Santiago, Chile, Sept. 1994.
[27] R. J. Bayardo: Efficiently mining long patterns from databases. SICMOD 1998: 85-93.
[28] R. Agarwal, C. Aggarwal, and V. V. V. Prasad: A Tree Projection Algorithm for Generation of Frequent Itemsets. ***Journal of Parallel and Distributed Computing***, 2000.
[29] M. J. Zaki: Scalable Algorithms for Association Mining. ***IEEE Transactions on Knowledge and Data Engineering,*** Vol. 12, No. 3, pp 372-390, May/June 2000.
[30] D. I. Lin and Z. M. Kedem: Pincer search: **A** new algorithm for discovering the maximum frequent sets. ***Proc. Of the 6th Int? Conference on Extending Database Technology (EDBT),*** Valencia, Spain, 1998.
[31] P. Shenoy, J. R. Haritsa, ***S.*** Sudarshan, G. Bhalotia, M. Bawa, and D. Shah: Turbo-charging Vertical Mining of Large Databases. SIGMOD Conference 2000: 22-33
[32] M. Holsheimer, M. L. Kersten, H. Mannila, and H.Toivonen: A Perspective on Databases and Data Mining. KDD 1995: 150-155
[33] A. Savasere, E. Omiecinski, and ***S.*** Navathe: An efficient algorithm for mining association rules in large databases. 21st VLDB Conference, 1995.
[34] V. Ganti, J. E. Gehrke, and R. Ramakrishnan: DEMON: Mining and Monitoring Evolving Data. ICDE 2000: 439-448.
[35] Doug Burdick, Manuel Calimlim and Johannes Gehrke. (2001). MAFIA: A Maximal Frequent Itemset Algorithm for Transactional Databases. *ieee*. p443-452.
[36] Asem Omari. (2010). A New Temporal Measure for Interesting Frequent Itemset Mining. *ieee*. p1-5.
[37] Wei Hou, Bingru Yang, Zhun Zhou and Chensheng Wu. (2008). An Adaptive Frequent Itemset Mining Algorithm for Data Stream with Concept Drifts. *ieee*. p1-4.
[38] CHEN Xiaoyun, HE Yanshan, CHEN Pengfei, MIAO Shengfa, SONG Weiguo, YUE Min. (2009). A Novel Parallel Frequent Itemset Mining Algorithm. *ieee*. p1-5.
[39] Guang-Peng Chen, Yu-Bin Yang, Yao Zhang. (2012). MapReduce-based Balanced Mining for Closed Frequent Itemset. *ieee*. p1-2.
[40] Ferenc Kov´acs,J´anos Ill´es. (2013). Frequent Itemset Mining on Hadoop. *ieee*. p1-5.
[41] CHEN Zi-Yang and LIU Guo-Hua. (2005). Quantitative Association Rules Mining Methods with Privacy-preserving. *ieee*.
[42] Yiping Ke ,James Cheng and Wilfred Ng. (2006). An Information-Theoretic Approach to Quantitative Association Rule Mining ⋆. *Springer LNCS*. p1-52.
[43] Om Prakash,Sunil Kumar,Anil Gupta. (2011). A Supplier Selection Criteria Using Boolean Association Rule Mining. *ieee*. p1-8.
[44] Fabrizio Angiulli, Giovambattista Ianni, and Luigi Palopoli. (1993). On the complexity of mining association rules. *ieee*. p1-8.
[45] Rolly Intan. (2003). AN ALGORITHM FOR GENERATING SINGLE DIMENSIONAL FUZZY ASSOCIATION RULE MINING . *ieee*. p1-6.
[46] Baoqing Jiang, Chong Han, Lingsheng Li. (1996). Mining Opened Frequent Itemsets to Generate Maximal Boolean Association Rules. *ieee*. p1-4.
[47] Z. G. Qu, X.X. Niu, J. Deng, C. McArdle, X.J. Wang. (2002). FREQUENT ITEMSET MINING OVER STREAM DATA: OVERVIEW. *ieee*. p1-6.

## AUTHOR DETAILS

**V. Aruna** has received the B.Tech degree (with Distinction) in Computer Science And Information Technology from the Sreenidhi Institute Of Science And Technology Affiliated to Jawaharlal Nehru Technological University of Hyderabad, in 2004, the M.Tech degree(with Distinction) in Computer Science Engineering from Sant Samarth Engineering College ,Bandlaguda Affiliated to Jawaharlal Nehru Technological University of Hyderabad, in 2010.And Currently pursing Phd (Computer Science And Engineering) in Jawaharlal Nehru Technological University of Hyderabad. Since 2004, she has been with Aurora's Scientific Technological and Reserach Academy, Bandlaguda, Hyderabad, where she is currently the

Head of the Department Computer Science and Engineering. Her main areas of research interest are Data mining And Warehousing, Computer Networks and Network security.

**V. Janaki.** received the Engineering degree in Electrical and Electronics Engineering from Andhra University in 1985, M.tech and in Computer Science and Engineering from REC Warangal in 1988 and Ph.D. degree in Computer Science and Engineering from Jawaharlal Nehru Technological University of Hyderabad in 2009. She is currently a Professor in the Department of Information Technology in Vaagdevi Engineering College, Warangal,. She has published papers in various National and International Journals and Conferences and organized various workshops and seminars. Her research interests include Network Security, Computer Networks.

**Dr. Avula Damodaram** joined as faculty of Computer Science & Engineering at JNTU, Hyderabad in the year 1989. In 2 decades of dedicated service, Dr. Damodaram performed distinguished services to the University as a Professor, Head of the Department, Vice Principal, Director of UGC-Academic Staff College, Director, School of Continuing & Distance Education and presently Director, University Academic Audit Cell. Dr. Damodaram has successfully guided 12 Ph.D. and 2 MS Scholars apart from myriad M. Tech projects. He is currently guiding 9 scholars for Ph. D and 1 scholar for MS. Dr. Damodaram is on the editorial board of 2 International Journals and a number of course materials. He is also the Chairman, Board of Studies for Computer Science and Engineering and Information Technology.

 