ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

A Complete Survey on Association Rule Mining and Its Improvement

Moushumi Sharma[1], Ajit Das[2], Nibedita Roy[3].

Pursuing M.SC(CS), Department of Computer Science and Technology, Bodoland University, Assam, India

Asst. Professor, Department of Computer Science and Technology, Bodoland University, Assam, India

Guest lecturer, Department of Computer Science and Technology, Bodoland University, Assam, India

ABSTRACT: In data mining, association rule learning is a popular and well researched method for discovering interesting relations between variables in large databases.Here we have classified Association rule mining in two ways, mining with candidate generation (Apriori Algorithm) and mining without candidate generation(FP-Tree). Further we have classified these two algorithms into different phases. Based on the limitations of these algorithms different researchers gave different ways to improve the efficiency of these algorithms. In this paper we present a survey of some research work carried by different researchers .We hope that it will provide a guideline for the researchers in interesting research directions that have yet to be explored.

KEYWORDS: Data mining, Association Rule Mining, Apriori Algorithm, FP Growth algorithm, Frequent Item-set mining.

I. INTRODUCTION

Data mining refers to the process of extracting or mining knowledge from large amounts of data. Association rule of data mining involves picking out the unknown inter-dependence of the data and finding out the rules between those items. Its aim is to extract interesting correlations, frequent patterns and association among set of items in the transaction database. It is used in all real life applications of business and industry. Association rules illustrate how frequently items are bayed together. E.g., an association rule beer => chips (80%) tells that four out of five customers that bought beer, they bought chips too. Such rules may become useful for decisions regarding product pricing, promotions, store outline and many others.

II. ASSOCIATION RULE

Association rule mining is the most important technique in the field of data mining. Association rules, first introduced in 1993 By Agrawal, are used to identify relationships among a set of items in databases. It consists of two procedures: First, finding the frequent item-set in the database using a minimum support and constructing the association rule from the frequent item-set with specified confidence. This mining is more applicable in the market basket analysis. This paper explains two important approaches, candidate generation approach and without candidate generation. This paper describes various methods improvements in the classical algorithms Apriori and FP growth‖ for frequent item-set generation. Apriori is used to find all frequent item-sets in a given database. The key idea of Apriori algorithm is to make multiple passes over the database. It employs an iterative approach known as a breadth-first search, through the search space ,where k-item-sets are used to explore (k+1) item-set. FP Growth Algorithm plays an essential role in association rule mining. It requires only two passes of processing. One pass is required for ordering and structuring frequent items, other pass is for inserting those frequent items in the tree. FP-tree as better performance than Apriori as reduce database scan. This paper introduces some new ways in which these algorithms can be improved.

III. MINING CANDIDATE GENERATION

Here we have discussed about the Apriori algorithm and different improvement algorithms on Apriori like AprioriTid, Apriori Hybrid, Dynamic Item-set Counting etc. To improve the efficiency of Apriori Algorithm different phases are given here.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

A. Minimize Candidate Item-Set

In Apriori algorithm, there is large number of candidate 2 item-sets and less tendency to determine support value. So, it takes lot of time to scan the database repeatedly and decreases the efficiency. To increase the efficiency of support calculation and reduce redundant pruning operation in Apriori algorithm the authors **Wanjun Yu, Xiaochun Wang, Fangyi Wang, Erkang Wang and Bowen Chen (2008) [4]** has given a new algorithm named as Reduced Apriori Algorithm with Tag (RAAT). To improve the above problem, RAAT uses Apriori-gen operation to form candidate 2-item-sets which results in diminishing the pruning operation. As a result, RAAT shortens the time and improves efficiency. The experimental results shown that RAAT performs well when it is compared with Apriori algorithm in a number of times. Authors **Huan Wu, Zhigang Lu, Lin Pan,RongshengXu and Wenbao Jiang (2009)** [5] proposed an improved Apriori algorithm (IAA) for association rule mining on the basis of analysis and study of previous efforts that researcher have applied. The IAA conquer the limitations of original apriori algorithm and introduces a new count based approach which is used to elicit the redundant candidate item-sets and uses generation record to reduce the scanning time of database. From the experimental results, it was proved that IAA is better than original Apriori algorithm because IAA counts each candidate item-sets once. To minimize the candidate generation, author **Sheila A. Abaya(2012)[1]** introduces a new way in which the Apriori algorithm can be improved. The modified algorithm introduces factors such as set size and set size frequency which in turn are being used to eliminate non significant candidate keys. The improved algorithm for Apriori takes for the set size which is the number of items per transaction and set size frequency which is the number of transactions that have at least "set size" items. In terms of execution time, the original apriori executes more time compared to the modified one. In terms of database passes, the modified apriori provides less database access compared with the original one that makes its execution faster. To mine efficient association rules using attributes and comparative analysis of various association rule algorithm, authors Ms **Shweta1 and Dr. KanwalGarg(2013)**[2] consider data (bank data) and tries to obtain the result using Weka a data mining tool. Here author consider three association rule algorithms: Apriori Association Rule, Predictive Apriori Association Rule and Tertius Association Rule. Author compares the result of these three algorithms and presents the result. According to the result obtained using data mining tool, author find that Apriori Association algorithm performs better than the Predictive Apriori Association Rule and Tertius Association Rule algorithms. Author **Y. Jia*et.al[26]*** proposed an improved algorithm based on combination of data division and dynamic item-set counting. The proposed algorithm has improved the two main problems which are faced by classical apriori algorithm. First is the repeatedly scanning of transactional database and second is the generation of large number of candidate sets. In data division, the transactional database is divided into n parts that don't intersect each other. After data division, dynamic item-sets counting are used to decide candidate item-sets before scanning database every time. So, the whole process needs only twice the entire database scan.

B. Based on Dense Data (too many transaction)

In order to find more valuable rules, the author **Jiao Yabing[6](2013)** proposes an improved algorithm of association rules, the classical Apriori algorithm. The algorithm includes two key processes: connecting step and pruning step. For dense database as large amounts of long forms occur, the efficiency of this algorithm is higher than Apriori. Finally, the improved algorithm is verified, the results show that the improved algorithm is reasonable and effective, can extract more value information. To improve the performance of apriori algorithm, authors **Jyoti Arora, NidhiBhalla and SanjeevRao[7](2013)**discussed a number of techniques. A review of four different association rule mining algorithms Apriori, AprioriTid, Apriorihybrid and tertius algorithms and their drawbacks which would be helpful to find new solution for the Problems found in these algorithms. In AprioriTid algorithm, the database is not used at all for counting the support of candidate item-sets after the first pass. Apriori performs better then Apriori-Tid in the initial passes but in later passes AprioriTid has better performance than Apriori. Dueto this reason they use another algorithm called Apriori Hybrid Algorithm. AprioriTertius algorithm finds the rule according to the confirmation measures. It uses first order logic representation.

C. Reduce cost and time consumed in the database scan

The authors **WEI-MIN MA and ZHU-PING LIU (2008)** [14], proposed two revised algorithms based on Apriori: AMS (Algorithm for mining stronger association rules) and AMLS (Algorithm for mining less strong association rules)

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

which focus on three aspects: minimum support, minimum confidence and minimum interest. These algorithms works in the form of matrix to decrease the scanning time of database. On the basis of comparison of classical Apriori algorithm with AMS and AMLS, it was proved that AMS and AMLS are better than classical Apriori algorithm. In order to reduce the time consumed to generate candidate iem-sets and scanning the database repeatedly in Classical Apriori algorithm, the authors **Yanfei Zhou, Wanggen Wan, Junwei Liu and Long Cai(2010)**[15], described an improved Apriori algorithm. This improved Apriori algorithm consists of three segments: First is decreasing number of judgements during the time of generating frequent candidate item-sets. Secondly pruning frequent item-sets. Finally, optimize the database. The improved Apriori algorithm was compared with classical Apriori algorithm on the basis on different support degree, different number of trading services and different number of items. From this comparison, it was proved that improved Apriori algorithm improves performance, increases efficiency, and reduces the redundant operation while producing frequent item-sets and strong association rules. Author **Shuo Yang et al (2012)** [11] proposed a theorem to improve the traditional Apriori Algorithm. The traditional Apriori Algorithm takes more time to scan the database and increases the complexity and decreases efficiency. The proposed algorithm decreases the database access on the basis of customer habits. For applying improved Apriori algorithm to E-commerce, there will be a need to develop a shopping site because when customers visit the shopping site the system will automatically find out their next purchasing goods that already available in their shopping basket. So, it will save time and increases the efficiency and provides more benefit. According to experimental results, it was shown that improved Apriori Algorithm when compared with traditional Apriori algorithm is more efficient. The authors **Jaishree Singh, Hari Ram, Dr.J.S.Sodhi (2013)** [10] have introduced a modified Apriori Algorithm called an improved AprioriAlgorithm(IAA) to conquer the limitations of classical Apriori Algorithm. The classical Apriori algorithm scans the database many times. If database contains ample number of records, it takes huge time to scan the database which results in increasing I/O cost. The improved Apriori Algorithm reduces the scanning time by eliminating the transactions containing irrelevant records. By comparing improved Apriori Algorithm with classical apriori algorithm, it was shown that improved Apriori Algorithm is better on the basis of efficiency and optimization. Based on the Apriori algorithm analysis and research, authors **JayshreeJhaand and LeenaRagha(2013)[3]** point out the main problems on the application Apriori algorithm in EDM(Educational Data Mining) and propose a new improved Apriori algorithm with a main motive of reducing time and number of scans required to identify the frequent item-set and association rules among education data using bottom up approach. The Major advantage of this algorithm is that it avoids comparison of currently chosen transaction with other transaction to mine the frequent item set if the total Support value or count of the other transactions on which comparison needs to be done is lesser than the chosen transaction. It also replaces user defined arbitrary minimum support threshold value of standard Apriori with functional model based on standard deviation which means that this algorithm can be well used by a non data mining expert. Based on Apriori algorithm, authors **Mohammed Al-Maolegi and BassamArkok [9](2014)** indicate the limitation of the original Apriori algorithm of wasting time for scanning the whole database searching on the frequent item-sets, and presents an improvement on Apriori. In this paper, the improved Apriori reduces the time consumed in transactions scanning for candidate item-sets by reducing the number of transactions to be scanned. The time consumed to generate candidate support count in their improved Apriori is less than the original Apriori; their improved Apriori reduces the time consuming by 67.38%.Authors **Maragatham G & Lakshmi M [20]** discuss the various advancements in data mining using the association rule mining. The role of association rules in temporal mining, utility mining, statistical mining, privacy preservation mining, particle swarm optimizations etc are reviewed. Generally adding time factor to association rule is called temporal mining. In temporal mining, the start and end of the valid time is added of each transaction and it is more useful and formative than basic association rule mining. Utility mining defines the usefulness of the item-sets with utility value. Privacy preserving preserve the sensitive data and useful rules extracted from the database. The concept of post processing and filtering out the less relevant rules can be done by statistical measure.

D. Save Memory Space

Author **R. Chang et.al [27](2011),** proposed an APRIORI-IMPROVE algorithm in which level L2 is directly generated from one scan over the database without generating candidate sets C1, L1 and C2. APRIORI-IMPROVE uses hash table and efficient horizontal data representation. APRIORIIMPROVE also optimized strategy of storage to save time & space.The performance of APRIORI-IMPROVE is higher as compared to apriori and fp-growth.**AuthorM. Patel et.al[25](2013)** proposed many algorithms to mine association rule that uses support and confidence as constraint. They proposed a method based on support value that increase the performance of Apriori

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

algorithm and minimizes the number of candidate generated and removed candidate at checkpoint which is infrequent which interns reduces storage and time required to calculate support of candidate.

E. Positive And Negative Association Rule

In this work authors **ShellyAhuja and GurpreetKaur(2014)[16]**present an efficient mining based on association rule generation by using predefined efficient algorithm, they find the positive and negative association rules. The objectives of their proposed method are generating relevant data base, an efficient algorithm for rule based classification and simulation for data mining, analysis of mining on the basis of precision, recall and f measure. In their work, the data set that they take is of Thyroid disease patient. They apply Apriori algorithm for mining their data set. Firstly, candidate generation phase comes after that number of occurrence of frequent item-sets is counted. If the support factor is greater than the decided threshold, positive and negative rules are generated for large item-set with support. By applying these rules classification of data set is done. As a result they get the better Precision, Recall, F-measure value. As a result dataset has been categorized in two classes positive and negative.

F. Drawbacks

Apriori Algorithm can be very slow and the bottleneck is candidate generation. For example if the transaction Database has 10^4 frequent 1-itemsets, they will generate 10^7 candidate 2-itemsets even after employing the downward closure. To compute those with support more than min sup, the database need to be scanned at every level. It needs (n+1) scans, where n is the length of the longest patter

IV. MINING WITHOUT CANDIDATE GENERATION

FP-Tree comes under the category of mining without candidate generation. Here we have discussed about FP-Tree (Frequent Pattern Tree). To improve the efficiency of FP-tree, different methods are given here.

A. Reduce Cost And Time Consumed In The Database Scan

Authors**BarışYıldız and BelginErgenç[19](2010)**explain the FP-Growth and the Matrix Apriori association rule mining algorithms that work without candidate generation. When the performances of the various algorithms are considered, they noticed that in constructing a matrix data structure, the Matrix Apriori takes more time in comparison to constructing the tree structure for the FP-Growth. On the other hand, during finding item-sets phase they discovered that the matrix data structure is considerably faster than the FP-Growth at finding frequent item-sets thus retrieving and presenting the results in a more efficient manner. Based on Apriori algorithm, the authors **Akshita Bhandari1, Ashutosh Gupta and Debasis Das(2014)[8]** indicates the wasting time and space for scanning the whole database of the original Apriori algorithm and presents an improvement on Apriori by reducing that wasted time depending on scanning by implementing a mathematical formula. The data structure used in this approach is the frequent pattern tree which can also be used to generate conditional patterns and suitable trees can be drawn for all the items. Their Algorithm can be used in the library for finding the book that is most frequently read and also used in the grocery shop by the shopkeeper for finding frequently sold item-sets as this takes lesser time and it's easy to find the items so that shopkeeper can make profit. Author **LiwenYue(2015)[18]** introduced some algorithms for frequent pattern mining, all of them were based on U-Aprior or tree structure UF-growth or improvement of them. They introduced the US-streaming algorithm and SUF-growth algorithm which were used of mining frequent pattern form uncertain data. The efficiency of US-streaming and SUF-growth algorithm was checked by setting the Minimum support threshold and size of database respectively. Experiment result shows that the improved algorithms for uncertain data have good efficiency in reducing memory and run time, especially used in complete frequent patterns.

B. Save Memory Space

Authors **Jiawei Han, Jian Pei and Yiwen Yin(2000)[13]**propose a novel frequent pattern tree structure, which is an extended prefix tree structure and develop an efficient FP-Tree based mining method, FP-growth, for mining the

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

complete set of frequent patterns by pattern fragment growth. Their improved FP-Tree avoids costly repeated database scans, avoid the costly generation of a large number of candidate sets and reduces the search space. Their performance study shows that the FP-growth method is efficient and scalable for mining both long and short frequent patterns and is faster than the Apriori algorithm and some recently reported new frequent pattern mining methods. **Most of the Incremental** rule mining methods are highly dependent on availability of main memory. If sufficient amount of main memory is not available, they fail to generate the results. Here authors **JyotiJadhav, LataRagha and Vijay Katkar (2012)[12]**presents a novel method for incremental discovery of frequent patterns using Main Memory database Management System to eliminate this drawback. Experimental results are provided to support the efficiency of proposed method. This method takes only one database scan and pass for processing frequent pattern. It also works efficiently in single as well as multiprocessing environment which gives better and faster performance than other existing algorithms. Author **Arpan Shah and Pratik A Patel (2014)[24]**explained fundamentals of frequent item-set mining. From the large variety of capable algorithms that have been established they compared the most important ones. They organize the algorithms and investigate their run time performance. The performance of algorithms reviewed in this paper based on support count, size of datasets, nature. Among all types of frequent item-set mining algorithm, FP growth is most capable and efficient technique to find frequent item-sets. FP growth technique is suitable for all kind of datasets. It constructed conditional structure to find relevant item-sets without candidate generation. FP growth is less memory consume so, the efficiency of the algorithm is effective. Authors**DeepakGarg and Hemant Sharma[21]**presents review of different frequent mining techniques including apriori based, partition based, DFS and hybrid, pattern based, SQL based and Incremental apriori based algorithms. A brief description of each technique has been provided. In the last, different frequent pattern mining techniques are compared based on various parameters of importance. Experimental results show that FP- Tree based approach achieves better performance by requiring only two database scans hence reducing the computational time. It takes less memory by representing large database in compact tree structure. In this paper authors **Sushila S. Shelke and Suhasini A. Itkar (2015)[22]**review the progress on techniques used in sequential pattern mining and distributed sequential pattern mining. This paper will explain the evaluation factors, the approach they are following and pros and cons of those algorithms. Finally they explain a little of demanding issues that need to be solved in future. Their studies in sequential pattern mining conclude that pattern growth approach is best suitable for further research effort in this region due to divide and conquer policy, no candidate generation and compressed database.

C. Positive and negative association rule

Sentiment analysis, a large majority of studies focus on identifying the polarity of a given text that is to automatically identify if a review about a certain topic is positive or negative. Authors **Ms. Prajacta Lobo and Prof. RajendraGawali (2016)**[23] have found this polarity by finding frequent item-set using proposed FP-growth and FIN method by making variation in Apriori. They compare the performance of both algorithms. The experimental result shows that FIN is more efficient in terms of memory consumption but more execution time taken compared to FP-growth. It also reduces the scanning of database and needs only two scanning of database

D. Mining Frequent Item

In this paper authors **Krutika. K .Jain and Anjali . B. Raut(2015)[17]** review the various trends of data mining and its relative applications. This paper is an attempt to use data mining as a tool used to find the hidden pattern of the frequently used item-sets. In the traditional association rules mining with FP-Trees and reduction technique minimum support and confidence threshold values are hard to be set for mining frequent item-sets without specific knowledge, users have difficulties in setting the support threshold to obtain their required results. The research part of this paper is this by changing the value of minimum confidence and gives different association rules. If the value of minimum confidence is high then rules filtered more accurately.

E. Drawbacks

FP-Tree may not fit in the main memory and it is expensive to build. In FP-tree, time is wasted specially if support threshold is high. Support can only be calculated once the entire dataset is added to the FP-Tree. At very low support, the number of frequent items became large and none of the algorithms were able to handle the large frequent sets gracefully.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

V. CONCLUSION

The problem of mining association rules has attracted lots of attention in the research community. In this paper, we have presented a survey of most recent work that has been done in Association rule based mining using Apriori algorithm and FP-Tree. The main intension of this paper is to review the progress on techniques used in Apriori and FB tree algorithm. After doing survey of various algorithms, we can make a conclusion that in improved Apriori algorithm the main focus is on generating less candidate sets which contains all frequent items within a reasonable amount of time. FP-Tree works better than Apriori algorithm in terms of database scan, cost and time. Hopefully, this short survey may provide a rough outline of the recent work and give people a general view of the field.

REFERENCES

[1]**Sheila A. Abaya** "Association Rule Mining Based On Apriori Algorithm In Minimizing Candidate Generation". International Journal of Scientific & Engineering Research Volume 3, Issue 7, July-2012 1 ISSN 2229-5518 IJSER © 2012 http://www .ijser.org

[2] Ms Shweta1 Dr. KanwalGarg"Mining Efficient Association Rules Through Apriori Algorithm Using Attributes and Comparative Analysis of Various Association Rule Algorithms" International journal of Advanced Research in Computer Science and Software Engineering, Kurukshetra University, India**, Volume 3, Issue 6, June 2013 ISSN: 2277 128X**

[3] Jayshree Jha1 and Leena Ragha2"Educational Data Mining using Improved Apriori Algorithm", International Journal of Information and Computation Technology. Department of Computer Engineering, RamaraoAdik Institute of Technology, Navi Mumbai, India. ISSN 0974-2239 Volume 3, Number 5 (2013), pp. 411-418.

[4] **Wanjun Yu, Xiaochun Wang and Fangyi Wang, Erkang Wang, Bowen Chen,** "The Research of Improved Apriori Algorithm for Mining Association Rules" 2008 11th IEEE International Conference on Communication Technology Proceedings, 978-1-4244-2251-7/08/$25.00 ©2008 IEEE.

[5] **Huan Wu, Zhigang Lu, Lin Pan, RongshengXu, Wenbao Jiang,** "An Improved Apriori-based Algorithm for Association Rules Mining" 2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery, 978-0-7695-3735-1/09 $25.00 © 2009 IEEE DOI10.1109/FSKD.2009.193.

[6] **Jiao Yabing** "Research of an Improved Apriori Algorithm in Data Mining Association Rules ".*International Journal of Computer and Communication Engineering, Vol. 2, No. 1, January 2013.*

[7] **Jyoti Arora1, Nidhi Bhalla2, Sanjeev Rao3**"A Review On Association Rule Mining Algorithms" International Journal of Innovative Research in Computer and Communication Engineering , Swami Vivekanand Institute of Engineering & Technology, Vol. 1, Issue 5, July 2013.

[8] **Akshita Bhandari1, Ashutosh Gupta2 and DebasisDas3 "**Improved Apriori Algorithm Using Frequent Pattern Tree For Real Time Applications". Student, Department of Computer Science and Engineering, NIIT University, Rajasthan, India. 2014

[9**] Mohammed Al-Maolegi1, Bassam Arkok2,** "An Improved Apriori Algorithm For Association Rules". Computer Science, Jordan University of Science and Technology, Irbid, Jordan. International Journal on Natural Language Computing (IJNLC) Vol. 3, No.1, February 2014 DOI : 10.5121/ijnlc.2014.3103 21

[10] **Jaishree Singh, Hari Ram andDr.J.S.Sodhi,** "Improving efficiency of Apriori algorithm using Transaction Reduction" International Journal of Scientific and Research Publications, Volume 3, Issue 1, January 2013 ISSN 2250-3153.

[11] **Shuo Yang,** "Research and Application of Improved Apriori Algorithm to Electronic Commerce" 2012 11th International Symposium on Distributed Computing and Applications to Business, Engineering & Science, 978-0-7695-4818-0/12 $26.00 © 2012 IEEE DOI 10.1109/DCABES.2012.51

[12] JyotiJadhav, LataRagha and Vijay Katkar 'Incremental Frequent Pattern Mining'. International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-1, Issue-6, August 2012

[13] **Jiawei Han, Jian Pei and Yiwen Yin,** "Mining Frequent Patterns without Candidate Generation" School of Computing Science,Simon Fraser University. SIGMOD'2000 Paper ID: 196

[14] **WEI-MIN MA and ZHU-PING LIU,** "TWO REVISED ALGORITHMS BASED ON APRIORI FOR MINING ASSOCIATION RULES", Proceedings of the Seventh International Conference on Machine Learning and Cybernetics, Kunming, 12-15 July 2008, -1-4244- 2096-4/08/$25.00 ©2008 IEEE.

[15] **Yanfei Zhou, Wanggen Wan, Junwei Liu and Long Cai,** "Mining Association Rules Based on an Improved Apriori.

[16] **Shelly Ahuja andGurpreet Kaur**, "Review of Association Rule Mining Using Apriori Algorithm". International Journal of Advanced Research in ComputerScience and Software Engineering .Chandigarh University,IndiaVolume 4, Issue 3, March 2014 ISSN: 2277 128X , 2014, IJARCSSE All Rights Reserved Page | 903

[17**] Krutika. K .Jain andAnjali . B. Raut**, "Review paper on finding Association rule using Apriori Algorithm in Data mining for finding frequent pattern" , International Journal of Engineering Research and General Science, HVPM College, Amravati, India ,Volume 3, Issue 1, January-February, 2015 ISSN 2091-2730 1008

[18] "Review of Algorithm for Mining Frequent Patterns from Uncertain Data". IJCSNS International Journal of Computer Science and Network Security, University of Yanshan, college of Information Science and Enginering, Qinhuangda, China. Vol.15 No.6, June 2015

[19] **BarışYıldızandBelginErgenç**"comparison of two association rule mining algorithms without candidate generation". International Conference on Artificial Intelligence and Applications(AIA 2010) Department of Computer Engineering,GulbahceKoyu, 35430 Urla

[20] Maragatham G &Lakshmi M "A Review On Association Rule Mining". Indian Journal of Computer Science and Engineering(IJCSE), Sathyabama University, Chennai, Tamil Nadu, India.

[21] **DeepakGarg and HemantSharma**,"Comparative Analysis of Various Approaches Used in Frequent Pattern Mining".International Journal of Advanced Computer Science and Applications, Thapar University, Patiala

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

[22] ***Sushila S. Shelke and Suhasini A. Itkar*** "A Review on Sequential Pattern Mining Algorithms". International Journal of Electrical, Electronics and Computer Engineering, 2015, PES Modern college of Engineering, Pune, (MH), India.
[23] "Sentiment Analysis using FP-Growth and FIN algorithm". International Journal of Research in Advent Technology, LokmanyaTilak College of Engineering, Mumbai University Vol.4, No.2, February 2016
[24] **Arpan Shah and Pratik A Patel** "A Collaborative Approach of Frequent Item Set Mining: A Survey" International Journal of Computer Applications. Parul Institute of Technology Gujarat, India Volume 107 □ No 8, December 2014
[25] Mihir R. Patel, Dipti P. Rana, and Rupa G. Mehta," FApriori: A Modified Apriori Algorithm Based on Checkpoint,"IEEE 2013.
[26] YuboJia, Guanghu Xia, Hongdan Fan, Qian Zhang and Xu Li, "An Improved Apriori Algorithm Based on Association Analysis," ICNDC 2012, 3rd IEEE International Conference, pp208-211.
[27] Rui Chang and Zhiyi Liu , " An Improved Apriori Algorithm," ICEOE 2011, IEEE International Conference, vol. 1, pp v1- 476 -v1-478.

BIOGRAPHY

Moushumi Sharma is a student of Computer Science And Information Technology Department, Bodoland University. Currently She is doing Master degree in Computer Science from Bodoland University, Assam. India. Her area of interest is in Data mining.

Ajit Das, Assistant Professor, Dept. of Computer Science and Technology, Bodoland University, Kokrajhar, Assam. He received Master in Computer Science(M.Sc in CS) in 2003 from Gauahati University, Assam, India and completed M.Phil from Periyar University Tamil Nadu, India in 2009.. His research interests are Speech Processing, Data mining etc.

Nibedita Roy, Guest lecturer and research scholar, Dept. of Computer Science and Technology, Bodoland University, Kokrajhar, Assam. She received Master degree in Mathematical Science in 2006 from Tezpur University, Assam, India and completed Master of Technology in Information technology from Tezpur university, Assam ,India in 2010.. Her research interests are Data mining, fuzzy set Theory, Image Processing etc.