# Review and Analysis of Apriori algorithm for Association Rule Mining

Shubhangi D. Patil
*Lecturer, Department of Information Technology,*
*Government Polytechnic, Jalgaon*

Dr. Ratnadeep R. Deshmukh
*Professor, Department of Computer and IT,*
*Dr. Babasaheb Ambedkar Marathwada University, Aurangabaad*

**Abstract- Data mining is a computerized technology that uses complicated algorithms to find relationships in large data bases Extensive growth of data gives the motivation to find meaningful patterns among the huge data. Apriori is a classic algorithm for frequent item set mining and association rule learning over transactional databases. Main idea of this algorithm is to find useful patterns between different set of data. The frequent item sets determined by Apriori can be used to determine association rules which highlight general trends in the database. It is a simple algorithm yet having many drawbacks. So many researchers have been contributed for the improvements of this algorithm. This paper does a survey on few good improved approaches of Apriori algorithm. This will be really very helpful for the upcoming researchers to find some new ideas from these approaches.**

**Keywords – Apriori algorithm, frequent itemset mining.**

## I. INTRODUCTION

Apriori [56] is a classic algorithm for frequent item set mining and association rule learning over transactional databases. It proceeds by identifying the frequent individual items in the database and extending them to larger and larger item sets as long as those item sets appear sufficiently often in the database. The frequent item sets determined by Apriori can be used to determine association which highlights general trends in the database: this has applications in domains such as market basket analysis

**Example 1** Consider the following database, where each row is a transaction and each cell is an individual item of the transaction:

| | | |
|---|---|---|
| Alpha | Beta | Epsilon |
| Alpha | Beta | Theta |
| Alpha | Beta | Epsilon |
| Alpha | Beta | Theta |

The association rules that can be determined from this database are the following: of sets with alpha also contain beta

1. 50% of sets with alpha, beta also have epsilon
2. 50% of sets with alpha, beta also have theta

We can also illustrate this through variety of examples

*Example 2*

Assume that a large supermarket tracks sales data by stock-keeping unit (SKU) for each item: each item, such as "butter" or "bread", is identified by a numerical SKU. The supermarket has a database of transactions where each transaction is a set of SKUs that were bought together.

Let the database of transactions consist of the sets {1,2,3,4}, {1,2}, {2,3,4}, {2,3}, {1,2,4}, {3,4}, and {2,4}. Apriori ia used to determine the frequent item sets of this database. To do so, we will say that an item set is frequent if it appears in at least 3 transactions of the database: the value 3 is the *support threshold*.

The first step of Apriori is to count up the number of occurrences, called the support, of each member item separately, by scanning the database a first time. We obtain the following result

| Item | Support |
|---|---|
| {1} | 3 |
| {2} | 6 |
| {3} | 4 |
| {4} | 5 |

All the itemsets of size 1 have a support of at least 3, so they are all frequent. The next step is to generate a list of all pairs of the frequent items:

| Item | Support |
|---|---|
| {1,2} | 3 |
| {1,3} | 1 |
| {1,4} | 2 |
| {2,3} | 3 |
| {2,4} | 4 |
| {3,4} | 3 |

The pairs {1,2}, {2,3}, {2,4}, and {3,4} all meet or exceed the minimum support of 3, so they are frequent. The pairs {1,3} and {1,4} are not. Now, because {1,3} and {1,4} are not frequent, any larger set which contains {1,3} or {1,4} cannot be frequent. In this way, we can *prune* sets: we will now look for frequent triples in the database, but we can already exclude all the triples that contain one of these two pairs:

| Item | Support |
|---|---|
| {2,3,4} | 2 |

In the example, there are no frequent triplets -- {2,3,4} is below the minimal threshold, and the other triplets were excluded because they were super sets of pairs that were already below the threshold.

We have thus determined the frequent sets of items in the database, and illustrated how some items were not counted because one of their subsets was already known to be below the threshold.

Apriori, while historically significant, suffers from a number of inefficiencies or trade-offs, which have spawned other algorithms.

1) It does multiple scan over the database to generate candidate set.

2) The number of database passes are equal to the max length of frequent item set.

3) For candidate generation process it takes more memory, space and time Candidate generation generates large numbers of subsets (the algorithm attempts to load up the candidate set with as many as possible before each scan).

Several improved algorithms have been proposed to conquer drawbacks of Apriori algorithm in several ways. Here we present six different approaches that face the common drawback. Rest of the improved apriori algorithms are summarized in table 1.

## II. IMPROVEMENT BASED ON DIVISION

Apriori is faced with two challenges. Maybe it needs scan transaction database repeatedly, increasing the I/O load of database.Probably it generates a large number of candidate sets, rasing the cost of time and storage space. To reduce the required amount of calculation and the times of scanning database, an improved algorithm combining Data

Division and Dynamic Itemsets Counting is proposed in [10], which can improve the performance meanwhile resolves the problems.

*Advantages*
1. It scans the database only twice.
2. It greatly reduces the I/O load.
3. It reduces the calculation and quantity of candidate sets and saves storage space to some extent.

## III. IMPROVEMENT FOR REDUCING THE NUMBER OF CANDIDATE ITEMSETS

Mihir R. Patel in [3] proposed a method that can be combined with Apriori algorithm and reduces storage required to store candidate and the execution time by reducing CPU time. CPU time is saved by reducing candidate sets size and time required to calculate the support of each candidate. The concept of checkpoint is introduced in [3] based on support value to reduce the execution time and overall storage space required to store candidate generated during scanning of dataset.

**Advantages**

The proposed method minimizes the number of candidate generated and removed candidate at checkpoint which is infrequent that reduces storage and time required to calculate support of candidate.

## IV. IMPROVEMENT BASED ON PARALLELIZATION

PApriori algorithm proposed by Ning Li [13] needs one kind of MapReduce job. The map function performs the procedure of counting each occurrence of potential candidates of size k and thus the map stage realizes the occurrences counting for all the potential candidates in a parallel way. Then, the reduce function performs the procedure of summing the occurrences counts. For each round of the iteration, such a job is carried out to implement the occurrences computing for potential candidates of size *k*

GPApriori [17] approach includes a new tree and vertical transaction list data structure and fine grain parallelization of the support counting algorithm

*Advantages*
1. The Proposed algorithm is actually more efficient as the database size is increased.
2. The execution time is reduced.
3. It improves the speed of the system

## V. IMPROVEMENT FOR REDUCING TIME AND SPACE COMPLEXITY

In [15], the HMT (HASH MAPPING TABLE) and HASH_TREE methodologies are used to
Optimize space complexity and time complexity. Using the HMT compressed Item sets, HASH_TREE can decentralize support count process. The result of experimental show that space complexity and time complexity of Apriori algorithm is efficiently reduced by using HMT and HASH_TREE**.**

Basic steps of Fast-Apriori algorithm [25] are same as those of Apriori algorithm, but the realization of Fast-Apriori is completely different from that of Apriori. The first step of Fast-Apriori algorithm is simply counting all frequencies of item sets containing one element, then decides the largest one-dimension frequent item sets. At the Kth step, there are three phases, firstly calling Fast-Apriori-Gen (Lk-1) function to generate candidate project set Ck, secondly calling Fast-Count-Support (Ck) function to count supports of elements in candidate project set Ck, finally counting one dimension frequent item sets K in candidate project set Ck.

*Advantages*
1. It can handle relatively large number of data.
2. Efficiency can be improved from space and time complexity.
3. It improve the efficiency of mining frequent itemsets.

## VI. IMPROVEMENT BASED ON SPECIFIC ITEMSET MINING

The algorithm based on the classic Apriori algorithm [21] greatly improves the real_time quality of and reduces the complexity. At the same time the algorithm uses XML as the operated database. It avoids multiple cycle operations in the general database, and enhances the operation efficiency. Algorithms directly delete model sets that users are

not interested in, and the results have met more constraints and have more pertinence, so the quality has also been greatly improved. In the detailed design process, authors program three functions, including k item sets generated by k-1 item sets, calculating and updating the support degree of candidate sets and the function judged whether an algorithm is completed. So in each database scanning process, it can produce different length frequent model sets and generate different length of the candidate frequent pattern sets in the same time so as to reduce the number of database scanning to gain all the frequent model sets which fulfills the restrictive conditions, decreases I/O costs and improves the efficiency of the mining.

The Apriori algorithm is a most influential one to excavate association rules. The basic idea of the algorithm is: identify all the frequent itemsets to get association rule. [22] presents the improved Apriori algorithm based on interested items, which mainly construct an ordered interested table and traverse it to excavate frequent itemsets quickly. The paper also by writing c# code achieves the improved algorithm. Confirmed by many experiments, this algorithm is better than traditional algorithms in time consuming

*Advantages*

1. It reduces the complexity.
2. Due to optimized algorithm, result's quality has greatly improved.
3. 3. It is better than traditional algorithm in time consuming.
4. 4. It reduces the frequency of scanning of database

## VII. CONCLUSION AND FUTURE SCOPE

It is very important to have a data mining algorithm with high efficiency because transaction database usually are very large. Various algorithms have been proposed for mining association rules but in every algorithm there are common drawbacks of multiple scans over the database. In this paper we have reviewed the various forms of improved apriori algorithm. Various authors have tried to improve the basic apriori algorithm based on parallelization, time efficiency, interested itemset mining, stopping condition etc. After surveying the various improved apriori algorithms we can conclude that the aim is to generate less candidate sets and yet get all frequent items. Survey concludes that many improvements are needed basically on pruning in Apriori to improve efficiency of algorithm. Considering frequency of items in database is also a good area to work on.

## ACKNOWLEDGMENT

The authors would like to welcome the reviewers for their critical and constructive comments and suggestions. The authors would like to thank the Authorities of Dr. Babasaheb Ambedkar Marathwada University, Aurangabad for providing the infrastructure to carry out the research

## REFERENCES

[1] Amimul Ehsan, Nagamma Patil, "Normalized Weighted and Reverse Weighted Correlation Based Apriori Algorithm", 978-1-4799-8792-4/15/2015 IEEE
[2] Tianpeng Han, Daimu Wang, "An Applied Research Based on Improved Association Rules", 2015 5th International Conference on Information Science and Technology (IC1S T) April 24-26, 2015, Changsha, Hunan, China.
[3] O.Jamsheela, Raju.G, "An Adaptive Method for Mining Frequent Itemsets Efficiently: An Improved Header Tree Method", 978-1-4799-8792-4/15/$31.00, 2015 IEEE.
[4] Wang Min, Zhang Hui, Yang Longxiang, Zhu Hongbo ,"A Two-dimensional Apriori Based User Behavior Prediction Algorithm In Mobile Social Environment" 978-1-4673-7687-7/15/$31.00, 2015 IEEE
[5] Yining Teng, Lanshan Zhang, Ye Tian, Xiang Li, "A Novel FAHP Based Book Recommendation Method by Fusing Apriori Rule Mining", 2015 International Conference on Intelligent Systems and Knowledge Engineering.
[6] Shaosong Yang, Guoyan Xu, Zhijian Wang, Fachao Zhou, "The Parallel Improved Apriori Algorithm Research Based on Spark", 2015 Ninth International Conference on Frontier of Computer Science and Technology
[7] Deepak Vidhate, Dr. Parag Kulkarni, "To improve Association Rule Mining using New Technique: Multilevel Relationship Algorithm towards Cooperative Learning", 2014 International Conference on Circuits, Systems, Communication and Information Technology Applications (CSCITA)
[8] Pel W Ang, Chunhong A N ,Lei W Ang, "An Improved Algorithm For Mining Association Rule In Relational Database", Proceedings of the 2014 International Conference on Machine Learning and Cybernetics, Lanzhou, 13-16 July, 2014.
[9] ZHOU Zhiping, WANG Jiefeng, "An Improved Matrix Sorting Index Association Rule Data Mining Algorithm", Proceedings of the 33rd Chinese Control Conference July 28-30, 2014, Nanjing, China.
[10] Zhao Yan Jun, Wei Ming Jun, Wang Jing, "Realization of Intrusion Detection System based on the Improved Data Mining Technology", IEEE 2013
[11] Johannes K. Chiang, Rui Han Yang, "Multidimensional Data Mining for Discover Association Rules in Various Granularities", IEEE 2013
[12] Mihir R. Patel, Dipti P. Rana, Rupa G. Mehta, "FApriori : A Modified Algorithm Based On Checkpoint", IEEE 2013
[13] Chao Tong, Peng Guo, "Data Mining With Improved Apriori Algorithm on Wind Generator Alarm Data" IEEE 2013

[14] Anand H.S. , Vinodchandra S.S, "Applying Correlation Threshold on Apriori Algorithm", IEEE 2013
[15] Mahesh Balaji, G Subrahmanya VRK Rao, "An Adaptive implementation Case Study of Apriori Algorithm for a Retail Scenario in a Cloud Environment", IEEE 2013
[16] Xingli Liu, Huali Liu, "An Improved Apriori Algorithm for Association Rules", IEEE 2013
[17] S. Suriya, S.P. Shantharajah,N "A Hybrid k-DCI and Apriori Algorithm for Mining Frequent Itemsets", IEEE 2013
[18] Yuntao Liu, Qing Liu, Danwen Zheng, Qingping Deng, Zangpeng Tan, Xiaoyang Jin, Wei Huang, Yaling Lei, Yi luo, "Application and Improvement Discussion about Apriori Algorithm of Association Rule Mining in Case Mining of Influenza Treated by Contemporary Famous Old Chinese Medicine", IEEE 2012
[19] Yubo Jia, Guanghu Xia, Hongdan Fan, Qian Zhang, Xu Li, "An Improved Apriori Algorithm Based on Association Analysis", IEEE 2012
[20] Li Dancheng, Zhang Ming, Zhou Shuangshuang, Zheng Chen, "A New Approach of Self Adaptive Discretization to Enhance the Apriori Quantitative Association Rule Mining", IEEE 2012
[21] Rupeng Luan, Sufen Sun, Junfeng Zhang,Feng Yu, Qian Zhang, "A Dynamic Improved Apriori Algorithm and Its Experiments in Web Log Mining", IEEE 2012
[22] Ning Li, Li Zeng, Qing He, Zhongzhi Shi, "Parallel Implementation of Apriori Algorithm Based on MapReduce", IEEE 2012
[23] Shuo Yang, "Research and Application of Improved Apriori Algorithm to Electronic Commerce", IEEE 2012
[24] Zhiyong Zeng, Hui Yang, Tao Feng, "Using HMT and HASH_TREE to Optimize Apriori Algorithm", IEEE 2011
[25] Yong Sen Jiang, Chu Xiang Chen, "The Data Mining of TCM Syndrome Dignostic Criteriaby the R_Apriori Algorithm", IEEE 2011
[26] Fan Zhang, Yan Zhang, Jason Bakos, "GPApriori :GPU- Accelerated Frequent Itemset Mining", IEEE 2011
[27] Masayuki Ohrui, Hiroaki Kikuchi, Masato Terada, Nur Rohaman Rosyid, "Apriori-Pre*fix*Span Hybrid Approach for Automated Detection of Botnet Coordinated Attacks", IEEE 2011
[28] Qiang Yang, Yanhong Hu, "Application of Improved Apriori Algorithm on Educational Information", IEEE 2011
[29] Li Min, Wang Chunyan, Yan Yuguang, "The Research of FP-growth method based on Apriori algorithm in MDSS", IEEE 2010
[30] Hongwei Xie, Xueli Yu, Juanli Li, Xiaobo Wu, "Study on Ontology- based Apriori Algorithm Applying to Emergency Decision System", IEEE 2007
[31] Libing Wu, Kui Gong, Fuliang Guo, Xiaohua Ge, Yilei Shan, "Research on Improved Apriori Algorithm Based on Interested Table", IEEE 2010
[32] Guofeng Wang, Xiu Yu, Dongbiao Peng, Yinhu Cui, Qiming Li, "Research of Data Mining Based on Apriori Algorithm in Cutting Database", IEEE 2010
[33] Chia Chu Chiang, Shen Lu, "Distributed Association Mining on Message Passing Systems", IEEE 2010
[34] Wang Hao Yu, Ji Xiao Juan, Xue Yun, Liu Xing, "Applying Fast Apriori Algorithm to Design Data Mining Engine", IEEE 2010
[35] Parvinder S. Sandhu, Dalvinder S. Dhaliwal, S. N. Panda, Atul Bisht, "An Improvement in Apriori Algorithm Using Profit and Quantity", IEEE 2010
[36] Yongge Shi, Yiqun Zhou, "An Improved Apriori Algorithm", IEEE 2010
[37] Gang Fang, Jiang Xiong, Cheng Sheng Tu, Ai Ping Luo, "An Algorithm of Mining Spatial Topology Association Rules Based on Apriori", IEEE 2010
[38] K.R.Suneetha, R. Krishnamoorti, "Advances Version of Apriori Algorithm", IEEE 2010
[39] Jean Daniel Cryans, Sylvie Ratte, Roger Champagne, "Adaptation of Apriori to MapReduce to Build a Warehouse of Relations Between Named Entities Across the Web", IEEE 2010
[40] Vaibhav Sharma, M.M. Sufyan Beg, "The probabilistic Approach to Apriori Algorithm", IEEE 2010
[41] Du Ping, Gao Yongping, "A New Improvement of Apriori Algorithm for Mining Association Rules", IEEE 2010
[42] Zalak V Vyas, Amit P. Ganatra, Dr. Y.P.Kosta, C.K. Bhesadadia, "Modified RAAT (Reduced Apriori Algorithm Using Tag) for Efficiency Improvement with EP(Emerging Patterns) and JEP(Jumping EP)", IEEE 2010
[43] Huang Chiung Fen, Tsai, Wen Chih, Chen An Pin, "Application of New Apriori Algorithm MDNC to Exchange Traded Fund", IEEE 2009
[44] Liu Yongge, Han Yahong, "Application of Apriori Algorithm in Oracle Bone Inscription Explication", IEEE 2008
[45] Huan Wu, Zhigang Lu, Lin Pan, Rongsheng Xu, "An Improved Apriori based Algorithm for Association Rules Mining", IEEE 2009
[46] Liu Jing, Qiu Chu, Lu Yongquan, Ji Haipeng, Wang Jintao, Li Nan, Gao Pengdong, Yu Wenhua, "An improved Apriori Algorithm for Early Warning of Equipment Failure", IEEE 2009
[47] Zhang Changsheng, Li Zhongyue, Zheng Dongsong, "An Improved Algorithm for Apriori", IEEE 2009
[48] Kan Jin, "An efficient algorithm for association mining", IEEE 2009
[49] Qihua Lan, Defu Zhang, Bo Wu, "A New Algorithm For Frequent Itemsets Mining Based on Apriori And FP-Tree", IEEE 2009
[50] Changsheng Zhang, Jing Ruan, "A Modified Apriori Algorithm With Its Application in Instituting Cross-Selling Strategies of the Retail Industry", IEEE 2009
[51] Shao Xiao Dong, "The Application of Improved 3D-Apriori Three-Dimensional Association Rules Algorithm in Reservoir Data mining", IEEE 2009
[52] Weidong Mao, "The Application of Apriori-Gen Algorithm on the Interaction Analysis of SNPs in Genetic Disease", IEEE 2008
[53] Sean Chester, Ian Sandler, Alex Thomo, "Scalable APRIORI- Based Frequent Pattern Discovery", IEEE 2009
[54] JianFeng Xu, Zhuo Wang, Lan Hu, ChauanHua Huang, "Rough_Apriori algorithm and The Application of an Aid System of Campus Student Major Selection", IEEE 2009
[55] Pan Weihao, Sun Jinguang, "Research for Parallel Apriori Algorithm based on MPI", IEEE 2009
[56] Sandeep Singh Rawat, Laxmi Rajamani, "Performance of Distributed Apriori Algorithm on a Computational Grid", IEEE 2009
[57] Yiwu Xie, Yutong Li, Chunli Wang, Mingyu Lu, "The optimization and Improvement of the Apriori Algorithm", IEEE 2008
[58] Wanjun Yu, Xiaochun Wang, "The Research of Improved Apriori Algorithm for Mining Association Rules", IEEE 2008
[59] Bo Wu, Defu Zhang, Qihua Lan, Jiemin Zheng, "An Efficient Frequent Patterns Mining Algorithm based on Apriori Algorithm and the FP-Tree Structure", IEEE 2008
[60] Sheng Chai, Jia Yang, Yang Cheng, "The Research of Improved Apriori Algorithm for Mining Association Rules", IEEE 2007

Table 1: Review of Various improved Apriori Algorithms

| Sr. No. | Paper Title | Author | Year | Technique | Advantages | Disadvantages |
|---|---|---|---|---|---|---|
| 1 | Normalized Weighted and Reverse Weighted Correlation Based Apriori Algorithm | Amimul Ehsan, Nagamma Patil | 2015 | Normalized Weighted and reverse weighted correlation (NWRWC) based Apriori algorithm | Some researches have proposed the method of applying weights According to the importance of the items but in these methods Many items with high supporting degree but low weight get Pruned. NWRWC based Apriori algorithm is proposed to deal With this situation by applying direct normalized weights as well As reverse normalized weights to the items. It further establishes The relevance between itemsets using weighted correlation Methods. | |
| 2 | An Applied Research Based on Improved Association Rules | Tianpeng Han, Daimu Wang | 2015 | CSP-Apriori Algorithm for processing transaction data with 'up' values And for mining frequent itemsets from the transactions. | | |
| 3 | An Adaptive Method for Mining Frequent Itemsets Efficiently: An Improved Header Tree Method | O.Jamsheela, Raju.G | 2015 | A new Binary Search Header Three (BSHT) and an Improved Header Tree mining (IHT-growth) to improve the performance Of the frequent pattern mining. | By applying The efficient binary search tree, the mining time drastically Reduced. The new transaction sorting method is also improves The performance of mining. | |
| 4 | A Two-dimensional Apriori Based User Behavior Prediction | Wang Min, Zhang Hui, Yang Longxiang, Zhu Hongbo | 2015 | An improved optimal Weighted fusion method | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Algorithm In Mobile Social Environment | | | based on effectiveness factor | | |
| 5 | A Novel FAHP Based Book Recommend ation Method by Fusing Apriori Rule Mining | Yining Teng, Lanshan Zhang, Ye Tian, Xiang Li, | 2015 | Fuzzy Analytical Hierarchy Process (FAHP) based method by fusing Apriori rule mining | In contrast to Conventional work, which only depends on few factors or could Not work on sparse data reliably, the proposed approach can Make efficient recommendation depend on different kinds of Factors. | |
| 6 | The Parallel Improved Apriori Algorithm Research Based on Spark | Shaosong Yang, Guoyan Xu, Zhijian Wang, Fachao Zhou | 2015 | Apriori Algorithm in multiple levels so called Multilevel Relationship Algorithm (MRA) | MRA Performs better than the Apriori algorithm towards Improvement of mining association rule. | Though this algorithm is Good to find the frequent item sets with minimum support it Does not provide with dependencies between different Frequent itemsets. |
| 7 | To improve Association Rule Mining using New Technique: Multilevel Relationship Learning Algorithm towards Cooperative | Deepak Vidhate, Dr. Parag Kulkarni | 2014 | Optimized method which searches the Candidate itemsets avoiding to scan the database repeatedly in Order to improve the efficiency of data mining | After the Optimization, weighted Apriori algorithm is more efficiently Because it only needs to scan the database once. As the Number of scanning the database reduces, the Apriori Algorithm running is more efficient than FP-growth under The same conditions. | |
| 8 | An Improved Algorithm For Mining Association Rule In Relational Database | Pel W Ang, Chunhong A N, ,Lei W Ang | 2014 | Improved Apriori Algorithm's parallel processing with spark support | Saves Dataset in memory mapped using data structures. | Although this Data structure can compress the original data set, whereas when The data set has a lot of transaction attributes and there are large Differences in latitude among the affairs, it will result in a Waste of memory resources |
| 9 | An Improved Matrix | ZHOU Zhiping, WANG | 2014 | Matrix-based and sorting index | Matrix algorithms Can improve the efficiency in computing frequency 2-itemset, | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Sorting Index Association Rule Data Mining Algorithm | Jiefeng | | association rules algorithm | but not delete non-frequency item set before calculation, not effectively Improved efficiency. | |
| 10 | Realization of Intrusion Detection System based on the Improved Data Mining Technology | Zhao Yan Jun, Wei Ming Jun, Wang Jing | 2013 | Improved Apriori Algorithm and K-Means Algorithm | | |
| 11 | Multidimensional Data Mining for Discover Association Rules in Various Granularities | Johannes K. Chiang, Rui Han Yang | 2013 | | 1.Greater Comprehensiveness and easy of use. 2.More efficient with limited scans and storage I/O operations. | |
| 12 | FApriori : A Modified Algorithm Based On Checkpoint | Mihir R. Patel, Dipti P. Rana, Rupa G. Mehta | 2013 | FApriori-The Proposed Algorithm introducing the concept of checkpoint based on support value. This concept is combined with Apriori Algorithm. | 1.The proposed method minimizes the number of candidate generated and removed candidate at checkpoint which is infrequent that reduces storage and time required to calculate support of candidate. | |
| 13 | Data Mining With Improved Apriori Algorithm on Wind Generator Alarm Data | Chao Tong, Peng Guo | 2013 | This paper improves Apriori Algorithm mainly by reducing the number of candidate itemsets. | 1.The Improved Apriori algorithm has higher execution efficiency than the traditional algorithm. | |
| 14 | Applying Correlation Threshold on Apriori Algorithm | Anand H.S. , Vinodchandra S.S. | 2013 | Correlation Threshold | 1.It reduces the Time complexity into O(n) from O(en). 2.The Proposed scheme creates more number of frequent itemsets in lesser time. | |
| 15 | An Adaptive implementation Case Study of Apriori Algorithm for a Retail Scenario in a Cloud Environment | Mahesh Balaji, G Subrahmanya VRK Rao | 2013 | Apriori Algorithm Based on Cloud Infrastructure | 1.The Apriori algorithm seems to indicate rapid mining of transaction data could happen in a very short time without any loss of information. | |

| | | | | | | |
|---|---|---|---|---|---|---|
| 16 | An Improved Apriori Algorithm for Association Rules | Xingli Liu, Huali Liu | 2013 | Improved Apriori | 1.The proposed algorithm improves the performance of the system. | |
| 17 | A Hybrid k-DCI and Apriori Algorithm for Mining Frequent Itemsets | S. Suriya, S.P. Shantharajah | 2013 | k-DCI algorithm is hybridized with Apriori algorithm | | |
| 18 | Application and Improvement Discussion about Apriori Algorithm of Association Rule Mining in Case Mining of Influenza Treated by Contemporary Famous Old Chinese Medicine | Yuntao Liu, Qing Liu, Danwen Zheng, Qingping Deng, Zangpeng Tan, Xiaoyang Jin, Wei Huang, Yaling Lei, Yi luo | 2012 | Improvement Based on the specific itemset, division, matrix, sampling, dynamic itemset counting | | |
| 19 | An Improved Apriori Algorithm Based on Association Analysis | Yubo Jia, Guanghu Xia, Hongdan Fan, Qian Zhang, Xu Li | 2012 | Improvement based on a combination of Data Division and Dynamic Itemset Counting | 1.It scans the database only twice.<br><br>2.It greatly reduces the I/O load.<br><br>3.It reduces the calculation and quantity of candidate sets and saves storage space to some extent. | |
| 20 | A New Approach of Self Adaptive Discretization to Enhance the Apriori Quantitative Association Rule Mining | Li Dancheng, Zhang Ming, Zhou Shuangshuang, Zheng Chen | 2012 | Approach to make Discretized Partitions in a Self Adaptive way to Enhance the Numeral Quantitative Result of Apriori Algorithm | 1.The proposed method can make better partitions automatically than traditional method. | |