Research Article

Zhenyi Zhao, Zhou Jian, Gurjot Singh Gaba, Roobaea Alroobaea, Mehedi Masud*, and Saeed Rubaiee

An improved association rule mining algorithm for large data

Abstract: The data with the advancement of information technology are increasing on daily basis. The data mining technique has been applied to various fields. The complexity and execution time are the major factors viewed in existing data mining techniques. With the rapid development of database technology, many data storage increases, and data mining technology has become more and more important and expanded to various fields in recent years. Association rule mining is the most active research technique of data mining. Data mining technology is used for potentially useful information extraction and knowledge from big data sets. The results demonstrate that the precision ratio of the presented technique is high comparable to other existing techniques with the same recall rate, i.e., the R-tree algorithm. The proposed technique by the mining effectively controls the noise data, and the precision rate is also kept very high, which indicates the highest accuracy of the technique. This article makes a systematic and detailed analysis of data mining technology by using the Apriori algorithm.

1 Introduction

After decades of research and practice, data mining technique has absorbed many disciplines results and formed a unique research branch. Undoubtedly, the research and application of data mining are very challenging. Data mining has to go through concept presentation, concept acceptance, extensive research and exploration, gradual application, and mass application stages like developing other new technologies. Most scholars believe that data mining research is still in the stage of extensive research and exploration from the current situation. On the one hand, the concept of data mining has been widely accepted. In theory, several challenging and prospective questions are being asked that are attracting more and more researchers. Since the concept of data mining was put forward in the 1980s, its economic value has emerged, and it has been advocated by many commercial manufacturers, forming a preliminary market [1].

Because the association rule find the relationship between items that cannot be found by traditional artificial intelligence and statistical methods, it has an important research value. At the same time, it

* **Corresponding author: Mehedi Masud,** Department of Computer Science, College of Computers and Information Technology, Taif University, Taif, Saudi Arabia, e-mail: mmasud@tu.edu.sa
Zhenyi Zhao, Zhou Jian: Wenzhou Vocational College of Science & Technology, Yong Jia, Wen Zhou, 325000, China
Gurjot Singh Gaba: School of Electronics & Electrical Engineering, Lovely Professional University, Phagwara 144411, India
Roobaea Alroobaea: Department of Computer Science, College of Computers and Information Technology, Taif University, Taif, Saudi Arabia
Saeed Rubaiee: Department of Industrial and Systems Engineering, College of Engineering, University of Jeddah, Jeddah, Saudi Arabia

satisfies people's urgent need to acquire knowledge from large-scale data storage. Currently, the research institutions of the world's famous universities and the major IT companies' research departments have invested much energy in their research and achieved many research results. It includes many advanced mining algorithms. Users who do not need to have advanced statistical knowledge and training can use it to dig out, including sequential patterns, classification, and so on the many types of knowledge. The system can run on various platforms, and many mainstream database systems (such as SQL-Server and Oracle) are closely combined. Simultaneously, online analysis and mining technology are also introduced so that the system can analyze advantages of data warehouses [1].

The computational procedure of expansive data sets' examples disclosure is included in data mining. The data are concentrated from the dataset, which further utilizes for the reasonable structure. Data mining is about taking care of issues by dissecting data in databases [2]. The organizations make proactive knowledge-driven decisions, and these tools predict future trends [3]. The knowledge discovery in database (KDD) is the basic step in the data mining techniques, and the data mining is referred to as KDD [4]. For since word alternatives, data mining utilized in KDD are care-of [5]. Figure 1 shows the data mining steps in knowledge discovery in database.
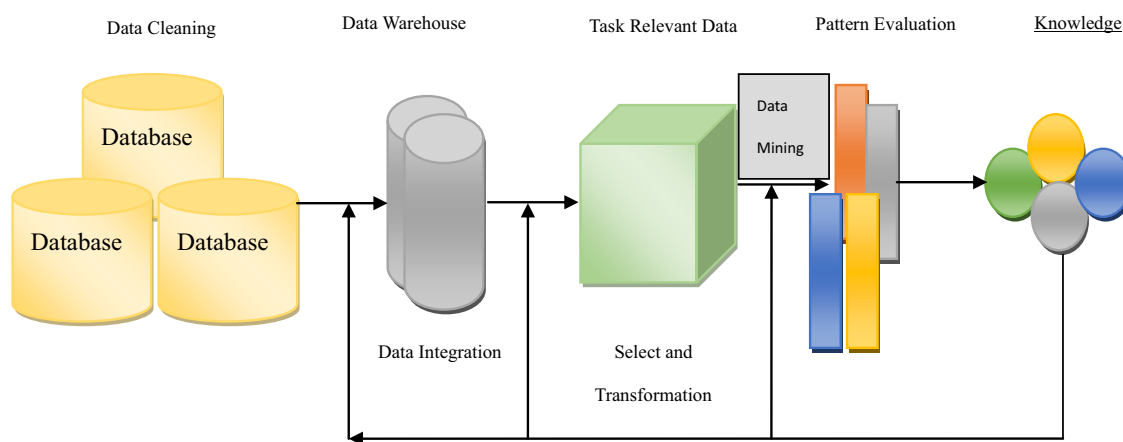

Figure 1: Basic diagram of data mining steps in database.

Data mining and association rules have attracted great attention in the information industry. Research institutions have carried out research and exploration on data mining technology.

The organization of this article is as follows. Section 2 provides a summary of the exhaustive literature survey followed by a methodology adopted in Section 3. A detailed discussion and analysis of the Apriori algorithm are given in Section 4. Section 5 provides detailed information that shows the improvement of the Apriori algorithm and details the mining data results. Finally, concluding remarks are provided in Section 6.

2 Literature review

Data mining can find out useful information that traditional analysis methods cannot find. Many famous universities in the world of the major research institutions and IT companies in the research department have spent a lot of energy to study and obtained many research results. For example, Stanford University developed a DMMiner mining system, which includes many advanced mining algorithms, mining the type of knowledge (AssociationRules) from association rules, and sequential patterns (sequence pattern) to find the classification of the drive (Discovery – Driver), and the system can run on multiple platforms, with many

mainstream database systems closely. IBM's Almaden lab's Quest project contains sequential patterns of association rules, classification, and clustering of the time series (TimeSeriesClustering) research. The representative products are DB2IntelligentMinerforData. Canada SimonFraser OBMiner was developed at the university. The system design aims to find the relationship between database and data mining integration based on the attribute-oriented concept of multistage found all kinds of knowledge. Many university research institutions and scholars made a great contribution to the development in this field. For example, Simon Fraster University in Canada and the University of Helsinki, Belgium, are famous in the world in data mining research. Moreover, there are numerous research works in this area. Cho et al. proposed the famous Apriori algorithm to improve the efficiency of mining association rules, and many new technologies were also generated [6].

Many authors have worked on data mining techniques. The traditional algorithms have been unable to meet data mining requirements in the aspect of efficiency [7]. The parallelization based on the Hadoop framework algorithm is realized. The existing data mining algorithms are highly complex, and the execution time is too long. In this article, the authors detail and analyze the association rules in the data mining technology and their merits and demerits [8]. The obtained results showed that the proposed algorithm is superior to the existing techniques. The authors in this article proposed a new data mining algorithm that is based on an association rule algorithm [9]. The K-means clustering algorithm is used for the clustering analysis of new mining results. The authors provide a relative study on a percentage of the most widely data mining algorithms that are used in commercial business and normal life [10]. Different systems, tools, and software's are used for relative data extraction from a specific group of data. The authors reviewed many data mining techniques [11] and recommend the products to the user based on the transaction history of other users who have the same characteristics as this user [12]. Hence, details such as age, gender, education, marital status, and salary are collected. So, data mining techniques such as clustering are required. The general Apriori Algorithm is used in the study. The authors explained the E-commerce businesses using tools of implicit rule algorithm in the data mining [13]. The experimental results obtained from the proposed techniques show that the processing time is improved.

2.1 Contribution

Different data mining existing technologies have some complexities in terms of time, computation, and efficiency. This article provides a systematic, in-depth, comprehensive research on the data mining technology. The data mining association rules are deeply analyzed by using the Apriori algorithm because it is an efficient algorithm from other state-of-the-art techniques. The results obtained from the improved Apriori mining algorithm show that it is not only simpler but also more efficient technique compared to other existing techniques.

3 Research on data mining technology

With the developed database techniques and the application of database management systems, the data storage in the world has increased rapidly. However, the current database system has not discovered the hidden knowledge behind the data and cannot predict the future trend of the development according to the data. The lack of technology and means for in-depth data analysis lead to the phenomenon of "rich data but poor knowledge" [14]. In the face of this challenge, data mining and knowledge discovery DM&KD technology emerged and developed rapidly.

3.1 Definition of data mining

The data forms knowledge like regard concepts, rules, patterns, and constraints. The hidden patterns that may exist in large databases are searched. The patterns and correlations between patterns are discovered by the data mining scans through large datasets. The data analysis and prediction are included in the data mining along with the collection and managing the data. The data are represented in quantitative, textual, or multimedia forms that can be performed on data mining. The data are examined by the variety of parameters used by the data mining applications [15]. Defining the problem, data preparation and exploration, building models, model exploration, and validation are the basic steps that can be defined. Researchers from different fields, especially in the database, artificial intelligence, machine learning, statistics, pattern recognition, data visualization, and other aspects of scholars and engineers, have been brought together to devote themselves to the emerging research field of data mining, forming a new technical hotspot [16].

However, compared with the traditional data analysis, the differences between data mining tools and traditional analysis tools are presented in Table 1.

Table 1: Differences between data mining tools and traditional analysis tools

	Data analysis tools	Data mining tool
Characteristics of tool	Retrospectives, validations	Predictive, discoverable
Analysis of the key	What has happened	Predict the future
Analysis purposes	List the largest customers from the most recent sales documents	Lock in future potential customers to reduce future sales costs
Data set size	The data dimension and the data in the dimension are small	The data dimension and the number of attributes in the dimension are huge
Start the way	Business manager, systems analyst, management consultant startup and control	Data with system startup, minimal personnel guidance
Technical conditions	Mature	Statistical analysis tools are being developed

3.2 Data mining technology

Data mining technology covers a wide range, mainly including database system, artificial intelligence, machine learning, data visualization, and other fields. There are also many technologies used in data mining. Figure 2 shows the structure diagram of the data mining technique. In data mining, we rarely use one tool or technique. For a given problem, the nature of the data itself affects how the technology is chosen, so we should use various techniques or tools to find the best model. The following is a brief introduction and analysis of the techniques often used in data mining.

The user interface in the starting provides the human computer interaction and communication. Model evaluation is an integral part of the model development process. It helps to find the best model that represents our data and how well the chosen model will work in the future. All the tools and software employed in the system are included in the engine to gain knowledge from the data warehouses. The data ware house compiled and organized all the data from the big data in the database. All the noisy data are removed from the database and to correct the inconsistencies in the data cleaning and selection process.
(1) Neural network

The artificial neural network image simulates human intuition thinking based on the characteristics of the biological neuron and neural network, by simplifying, inducing, and summarizing a class of parallel processing network. Neural networks for large-scale and complex problems containing hundreds of interactions between the independent variables can be effectively used for model, and hence, people are very interested in the neural network [17]. It can be used to solve classification
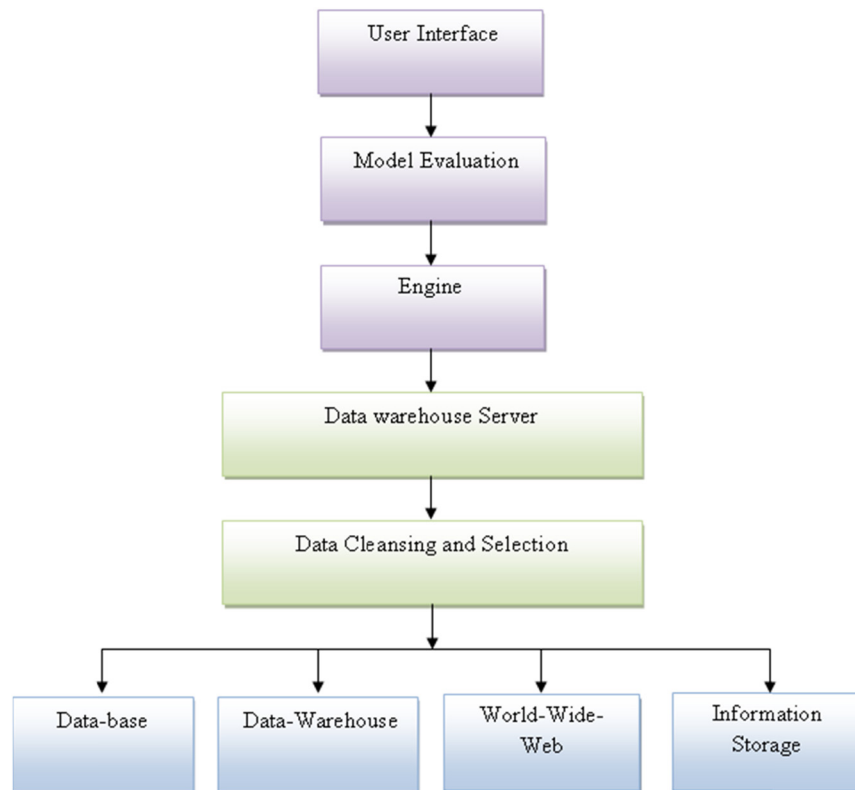
Figure 2: Technical structure of data mining.

problems in data mining (output variables are discrete) and regression problems (output variables are continuous).

(2) The decision tree

The data records are classified by the tree structure, and a record set is represented by a leaf node under a certain condition, and branches of the tree are established according to different values of record fields [18]. A decision tree is a way of deriving a class of values from a set of rules.

(3) Genetic algorithm

Genetic algorithms, by themselves, are not utilized to discover patterns but used to guide the learning process. Genetic technique guide is followed for finding good models and pattern of biological evolution, passing on their characteristics from generation to generation until they find the best model [19]. The information inherited is called a gene, which contains the parameters of the model established.

(4) Visualization technology

Data visualization techniques are often used in conjunction with other data mining techniques to analyze data effectively, and their importance cannot be underestimated [20]. For example, the multi-dimensional data in the database is changed into a variety of graphics, which play a great role in revealing the distribution of data. Visualization of the data mining process and man–machine interaction can improve the data mining effectiveness [21]. There are rough sets, regression analysis, discrimination analysis, and other techniques in addition to the aforementioned techniques.

3.3 The flow of data mining

The process of data mining includes problem definition, data preparation, data mining, result analysis, and knowledge application, as shown in Figure 3.
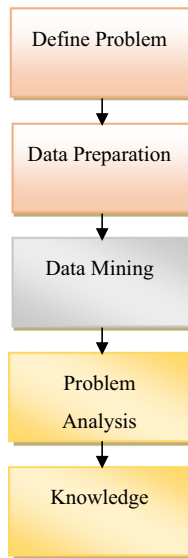
Figure 3: Data-mining flowchart.

First, the problem that to be solved is defined and then, the data relevant are decided and retrieved from the data collection for the analysis [22]. The transformation of the selected data into the forms of mining procedure is done. Data mining is a very important and a major step in which extract patterns are potentially extracted by clever techniques. The problem is then analyzed correctly [23]. The representation of the discovered knowledge is presented visually to the user, which is the final step for knowledge representation.

4 Classical association rule mining algorithm

Association mining algorithm is the active research area of data mining. The following algorithms, such as Apriori and partition, are briefly introduced.

4.1 Apriori algorithm

In 1994, Agrawal et al. designed a fundamental technique, Apriori, which is the influential and classical algorithm for mining one-dimensional, single-layer, and Boolean association rules.
(1) Design idea of Apriori algorithm

This algorithm is based on the idea of a two-stage frequent item set to obtain the method of finding the frequent item set:
1. Frequent items with a growth rate of 1 are recorded as L[1].
2. The candidate item set C[k+L] is generated, which is L[k] based, and the candidate's subsets are required to be frequent item sets.
3. The transaction D database is scanned, and each candidate item set is calculated, if it is greater than minsup, then attach to L[k+1].
4. If L[k+L] is a void set, it ends, and the desired outcome is L[1] union L[2] union. Otherwise, continue by turning 2.

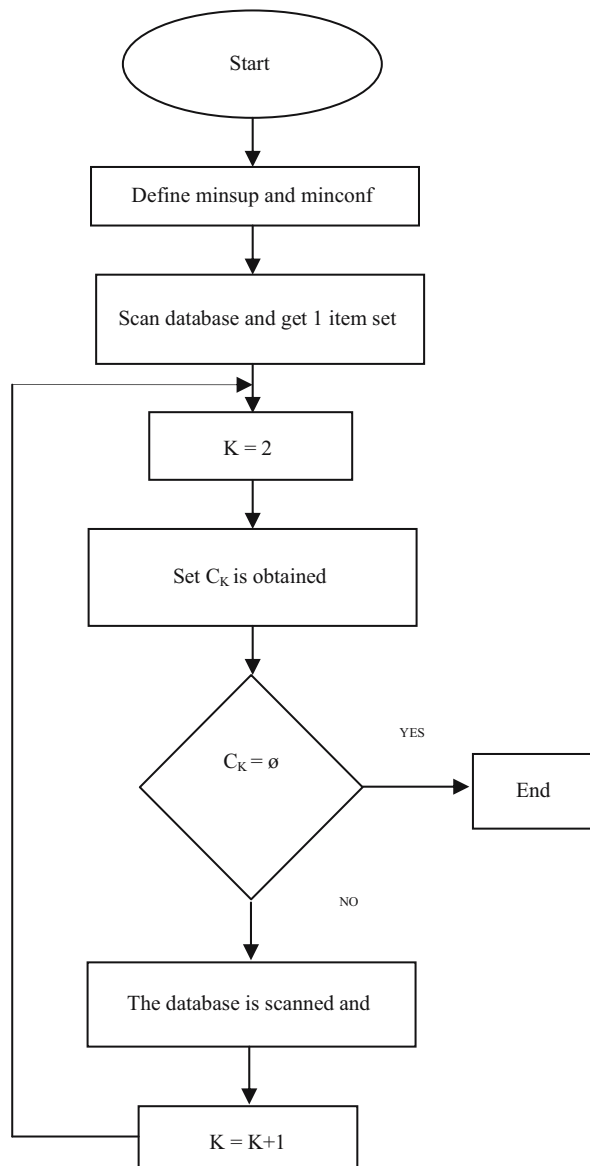Figure 4 provides the flow chart of the Apriori algorithm.

Figure 4: Flow chart of the Apriori algorithm.

(2) Apriori algorithm description

A test method is utilized by the Apriori algorithm called layered iteration, k-item set for exploring (k+1) [24]. First, the set L_1 of frequent 1-item sets is found. The set L_2 is found by the L_1 of frequent two-item sets, and L_2 is used to find the set L_3, and so on. A database scan is required for finding each L_k [25].

By definition, if the minimum support threshold is not met by the item set I, then I is not frequent, that is, support (I). If item A is added to I, the resulting item set (that is, IUA) cannot happen more commonly compared to I. Therefore, IUA is not frequent, that is, support (I).

The special classification called anti-monotone means that if a set is not passed the test, then all the supersets cannot pass the same test. It is called anti-monotone because it is monotone in the sense of not passing the test.

The Apriori algorithm is described as follows:

```
Input: Transaction database D.
Minimum support threshold minsup.
Output: frequent item set b in D.
    Begin
    L₁= find_frequent_1-itemsets(D):
    For (k =2 ; Lₖ₋₁≠φ ; k++) {
    Cₖ= apriori_gen(Lₖ₋₁,minsup);//New candidate
    For every transaction t∈D {
    Candidates included in Cₜ= subset(Cₜ,t) ;/ /t
    For each candidate c∈Ct
       C.count++;
    }

    Lₖ={c∈ Cₖ| c.count > minsup}

    }
    Return L=∪ₖLₖ
    End
```

Among them, apriori_gen is a frequent item set and L_{k-1} is connection candidate item set of C_k generation. The specific description process is as follows:

```
Procedure apriori_gen (Lₖ₋₁, minsup)
{
    For each itemset 1₁∈ Lₖ₋₁
    For each itemset 1₂∈ Lₖ₋₁
        if(1₁[1]= 1₂[1])∧(1₁[2]=1₂[2]) ∧...∧(1₁[k-2]= 1₂[k-2])∧  (1₁[k-1]<1₂[k-1])
        Then |
        {c=1₁.1₂; /*join step: generate candidates*/
           If has_infrequent_subset(c,Lₖ₋₁) then
               Delete c; /* prune step: remove unfruitful candidate*/
           Else add c to Cₖ
        }
        Return Cₖ; }
```

The frequent item sets must be frequent according to the Apriori property, and the layer-by-layer search technique is utilized by this algorithm. Given k-item sets, check (k–1) – subsets whether they are frequent. The frequent item sets must be frequent according to the Apriori property and layer-by-layer search technique is utilized.

```
        {
        For each (k-1)-subsets of c do
        If s ∉ Lₖ₋₁
        Then return TRUE
        Else return FALSEₖ }
```

4.2 Key technologies of Apriori algorithm

The Apriori property is applied in the L_{k-1} finding process of the algorithm, which is composed of connection and pruning [26].

(1) **Connection step:** L_k is found by connecting L_{k-1} with itself for k-item candidate set generation. The item sets are referred to as C_k [27]. Let l_1 and l_2 be the item set in L_{k-1}. The notation $L_i[j]$ represents the JTH item of L_i and, for convenience, assumes that the items in the item set are arranged in the lexicographical order [28].

(2) **Pruning step:** All k-frequent item sets are included in C_k and C_k is the superset of L_k, which may or may not be frequent [29]. To compress C_k, the Apriori property is utilized: any infrequent (k–1)-item set cannot be a subset of the frequent k-item set. The candidate is not frequent and can therefore be deleted from C_k [30].

5 Results and discussion

The features and the results obtained by the presented technique are discussed briefly in this section. The comparative analysis of the proposed technique with the state-of-the art techniques is also performed and discussed.

5.1 Features of Apriori algorithm

The algorithm explores the (k+1) item set according to the k-item set using a method called layer-by-layer iteration of candidate generation tests.

(1) Apriori algorithm is a hierarchical iterative algorithm

First, a set of frequent 1-item sets is found, which is denoted as L_1, then L_1 gets L_2, L_2 gets L_3, and so on, until the frequent k-item set cannot be found [31]. Apriori algorithm mining produces all frequent items with no less than minimum support of minsup.

(2) Data are organized in a transactional manner

The association is that the data are organized in the form of {Id,Item}, that is, {trans. number, Item set}.

(3) Pruning method was adopted

Using the property of frequent item set to optimize the search, because this optimization is the first in the algorithm, is called Apriori optimization [32,33]. Apriori optimization is essentially realized by pruning the candidate frequent item sets.

(4) Mining association rules applicable to transaction database.

5.2 Improvement of Apriori algorithm

To reduce the impact of existing problems in Apriori algorithm and improve the effectiveness of Apriori, many scholars have conducted a lot of research based on it and proposed some improved algorithms. These improved algorithms based on Apriori are usually called Apriori-like algorithms. The following is an introduction to several typical improvement methods [34].

(1) Hashing-based Optimization method

The optimization method based on hashing is utilized to compress the size of the (k ≥ 2) set of the candidate k-item set C_k. When the transaction database is scan, produced by the candidate k – itemsets, at the same time, it produce each transaction (k+l) subset and increase thecount barrels in the next

candidate item sets (k+1) [35–38]. This technique is particularly effective when k = 2. The key is to construct a valid hash function.
(2) Optimization method based on transaction compression

The transaction-based optimization method reduces the scanned transaction database size by reduction of unnecessary transactions, so as to improve the efficiency of mining [33]. We can delete these transactions because they are no longer needed when scanning the database to produce (k+1) item sets [39,40].
(3) Dynamic item set counting-based optimization method

The dynamic set counting-based optimization method divides the database into blocks. The algorithm can add a new set of waiting options at any starting point [40–43]. This technique dynamically evaluates the support of all item sets that have been counted. This algorithm has fewer times of scanning database than Apriori algorithm.

5.3 Data mining result verification in parallel

Data mining results are obtained and demonstrated by comparing serial and parallel programs of data mining. The parallel program is considered reliable if the results of the data parallel program are consistent with the serial data program. The data mining results of serial and parallel for 150 and 250 M file are tabulated in Tables 2 and 3, respectively. The FIM used in the table represents the frequent item sets.

Table 2: Data mining results of serial and parallel for 150 M files

Algorithm	Serial	Parallel
FIM1	21	21
FIM2	67	67
FIM3	45	45
FIM4	28	28

Table 3: Data mining results of serial and parallel for 250 M files

Algorithm	Serial	Parallel
FIM1	19	19
FIM2	67	67
FIM3	45	45
FIM4	28	28

By the parallel algorithms, there is a consistency in the serial algorithm with the varying item sets. The results from one sets to four sets are compared, but these are all consistent. The parallel technique is better in terms of reliability and accuracy and the frequent item sets are excavated accurately in which minimum support is satisfied. From the results, no such advantage of mining efficiency of the parallel algorithm was observed because of the work schedule overhead. The parallel algorithms is advantageous as it gradually emerges and the use less mining time than the serial algorithm. The proposed algorithm is also utilized for small-capacity database, and the time, the acceleration ratio, and the speed are presented in Table 4.

The parameters obtained by the presented technique is highly controlled and accurate. There system is highly reliable in terms of time, speed, and acceleration.

Table 4: The mining time, mining acceleration ratio, and mining speed are obtained

	Proposed technique
Mining time (s)	40.34
Mining acceleration	15.76
Mining speed	4.98

The presented method is compared with the existing techniques in terms of accuracy at the level of recall rate. The obtained results are tabulated in Table 5 and graphically represented in Figure 5. Graphical representation gives the better visualization of the values, and better analysis is done by the graphical form of values.

Table 5: Comparison of the proposed technique in terms of accuracy at the level of recall rate

Recall ratio (%)	R-tree (%)	Dynamic (%)	The proposed technique (%)
10	80.7	81.1	82.4
20	62.2	65.3	68.2
30	50.4	52.6	55.3
40	32.6	41.5	44.7
50	27.8	36.2	38.9
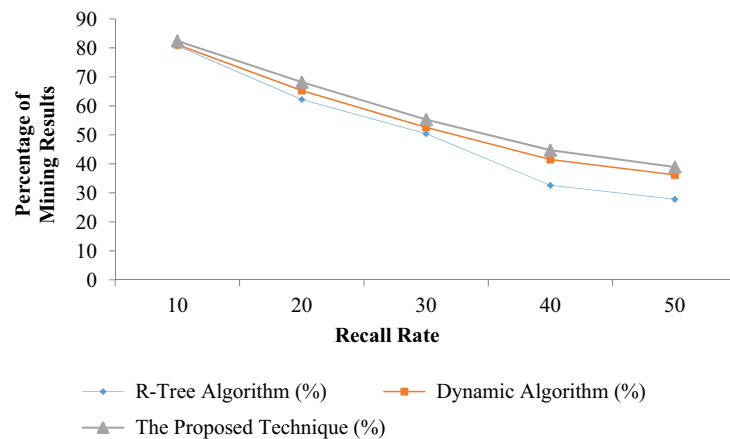
Figure 5: Comparison of the proposed technique in terms of accuracy at the level of recall rate.

It is clear from the comparison results that the precision ratio of the presented technique is high compared to other existing techniques with the same recall rate, i.e., R-tree algorithm. The noise data are effectively controlled by the proposed technique by controlling the mining time, speed, and the acceleration. The precision rate is also kept high, which indicates the highest accuracy of the technique.

6 Conclusion

The mining of association rules in large databases always requires more resources such as memory and CPU and expensive I/O costs, so improving efficiency is a work with high application value. The Apriori algorithm for association rules and the improved Apriori mining algorithm are further concluded that the algorithm is not only simple but also greatly reduces the number of candidate frequent item sets and

has the advantages of fast search speed, which not only saves the calculation cost but also improves the efficiency of the algorithm. The results obtained from the improved Apriori mining algorithm show that it is not only simpler but also more efficient technique compared to the existing techniques.

Conflict of interest: Authors state no conflict of interest.

References

[1]   Balaji BV, Rao VV. Improved classification based association rule mining. Int J Adv Res Comput Commun Eng. 2013;2(5):2211–21.
[2]   Hand DJ. Data mining: statistics and more? Am Stat. 1998;52(2):112–8.
[3]   De Clercq PA, Hasman A, Blom JA, Korsten HH. Design and implementation of a framework to support the development of clinical guidelines. Int J Med Inform. 2001;64(2–3):285–318.
[4]   Miller HJ, Han J. (Eds.). Geographic data mining and knowledge discovery. CRC press; 2009.
[5]   Bose I, Mahapatra RK. Business data mining – a machine learning perspective. Inf Manag. 2001;39(3):211–25.
[6]   Cho YH, Kim JK, Kim SH. A personalized recommender system based on web usage mining and decision tree induction. Expert Syst Appl. 2002;23(3):329–42.
[7]   Fu C, Wang X, Zhang L, Qiao L. Mining algorithm for association rules in big data based on Hadoop. AIP Conference Proceedings. Vol. 1955, no. 1. AIP Publishing LLC; 2018 April. p. 040035.
[8]   Xu Y. Research of association rules algorithm in data mining. Int J Datab Theor Appl. 2016;9(6):119–30.
[9]   Zhang G, Liu C, Men T. Research on data mining technology based on association rules algorithm. 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC). IEEE; 2019 May. p. 526–30.
[10]   Baker RSJD. Data mining for education. Int Encycl Educ. 2010;7(3):112–8.
[11]   Altaf W, Shahbaz M, Guergachi A. Applications of association rule mining in health informatics: a survey. Artif Intell Rev. 2017;47(3):313–40.
[12]   Yabing J. Research of an improved apriori algorithm in data mining association rules. Int J Comput Commun Eng. 2013;2(1):25.
[13]   Jha J, Ragha L. Educational data mining using improved apriori algorithm. Int J Inf Comput Technol. 2013;3(5):411–8.
[14]   Liu H, Gegov A, Cocea M. Rule based networks: an efficient and interpretable representation of computational models. J Artif Intell Soft Comput Res. 2017;7.
[15]   Chandaka Babi D, Rao MV, Rao VV. Study of association rule mining for discovery of frequent item sets on big data sets. Int J Appl Eng Res. 2017;12(22):12169–75.
[16]   Hussain S, Dahan NA, Ba-Alwib FM, Ribata N. Educational data mining and analysis of students' academic performance using WEKA. Indones J Electr Eng Comput Sci. 2018;9(2):447–59.
[17]   Masud M. Collaborative e-learning systems using semantic data interoperability. Comput Hum Behav. 2016;61:127–35.
[18]   Ramasamy S, Nirmala K. Disease prediction in data mining using association rule mining and keyword based clustering algorithms. Int J Comput Appl. 2020;42(1):1–8.
[19]   Liu H, Gegov A, Cocea M. Network based rule representation for knowledge discovery and predictive modelling. In 2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). IEEE; 2015 August. p. 1–8.
[20]   García E, Romero C, Ventura S, De Castro C. A collaborative educational association rule mining tool. Internet High Educ. 2011;14(2):77–88.
[21]   Nicholson S. The basis for bibliomining: Frameworks for bringing together usage-based data mining and bibliometrics through data warehousing in digital library services. Inf Process Manag. 2006;42(3):785–804.
[22]   Shi GR, Yang XS. Optimization and data mining for fracture prediction in geosciences. Proc Comput Sci. 2010;1(1):1359–66.
[23]   Liu Y. Research on association rules mining algorithm based on large data. Inf Technol. 2017;40(1):192–4.
[24]   Roobaea A. An empirical combination of machine learning models to enhance author profiling performance. Int J Adv Trends Comput Sci Eng. 2020;9:2.
[25]   Sharma A, Kumar R. A framework for pre-computated multi-constrained quickest qos path algorithm. J Telecommun Electron Computer Eng. 2017;9(3–6):73–7.
[26]   Feng H, Liao R, Liu F, Wang Y, Zhu X. Optimization algorithm improvement of association rule mining based on particle swarm optimization. In 2018 10th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA). IEEE; 2018. p. 524–9
[27]   Rizvi SJ, Haritsa JR. Maintaining data privacy in association rule mining. VLDB'02: Proceedings of the 28th International Conference on Very Large Databases. Morgan Kaufmann; 2002 January. p. 682–93.

[28] Afzali GA, Mohammadi S. Privacy preserving big data mining: association rule hiding using fuzzy logic approach. IET Inf Secur. 2017;12(1):15–24.

[29] Guangpeng L, Peng C. Research on reliability modeling of cnc system based on association rule mining. Proc Manuf. 2017;11:1162–9.

[30] Pannu A. Artificial intelligence and its application in different areas. Artif Intell. 2015;4(10):79–84.

[31] Hamzaoui A, Malluhi Q, Clifton C, Riley R. Association rule mining on fragmented database. Data privacy management, autonomous spontaneous security, and security assurance. Cham: Springer; 2014. p. 335–42.

[32] Jiang P, Liu XS. Big data mining yields novel insights on cancer. Nat Genet. 2015;47(2):103–4.

[33] Weng CH. Identifying association rules of specific later-marketed products. Appl Soft Comput. 2016;38:518–29.

[34] Kaushik M, Gupta SH, Balyan V. Power optimization of invivo sensor node operating at terahertz band using PSO. Optik. 2020;202:163530.

[35] Kaushik M, Gupta SH, Balyan V. Evaluating threshold distance by using eigen values and analyzing its impact on the performance of WBAN. In 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN). IEEE; 2019 March. p. 864–7.

[36] Masud M, Gaba GS, Alqahtani S, Muhammad G, Gupta BB, Kumar P, et al. A Lightweight and robust secure key establishment protocol for internet of medical things in COVID-19 patients care. IEEE Internet Things J. 2020;1–9.

[37] Masud M, Alazab M, Choudhary K, Gaba GS. 3P-SAKE: privacy-preserving and physically secured authenticated key establishment protocol for wireless industrial networks. Comput Commun. 2021;82–90.

[38] Rathee G, Sharma A, Iqbal R, Aloqaily M, Jaglan N, Kumar R. A blockchain framework for securing connected and autonomous vehicles. Sensors. 2019;19(14):3165.

[39] Masud M, Gaba GS, Choudhary K, Alroobaea R, Hossain MS. A robust and lightweight secure access scheme for cloud based E-healthcare services. Peer-to-Peer Netw Appl. 2021;1–15. doi: 10.1007/s12083-021-01162-x

[40] Poongodi M, Sharma A, Vijayakumar V, Bhardwaj V, Parkash Sharma A, Iqbal R, et al. Prediction of the price of Ethereum blockchain cryptocurrency in an industrial finance system. Comput Electr Eng. 2020;81:106527.

[41] Gaba GS, Kumar G, Monga H, Kim TH, Kumar P. Robust and lightweight mutual authentication scheme in distributed smart environments. IEEE Access. 2020;8:69722–33.

[42] Rathee G, Sharma A, Kumar R, Ahmad F, Iqbal R. A trust management scheme to secure mobile information centric networks. Comput Commun. 2020;151:66–75.

[43] Gaba GS, Kumar G, Monga H, Kim TH, Liyanage M, Kumar P. Robust and lightweight key exchange (LKE) protocol for industry 4.0. IEEE Access. 2020;8:132808–24.