World of Computer Science and Information Technology Journal (WCSIT)
ISSN: 2221-0741
Vol. 5, No. 11, 165-171, 2015

An Empirical Analysis of Multiple Level Association Rules Mining Method for Feature Extraction

Ruchika Yadav
Research Scholar,
Department of Computer Science and Application,
Kurukshetra University, Kurukshetra
Haryana, India

Dr. Kanwal Garg
Assistant Professor,
Department of Computer Science and Application,
Kurukshetra University, Kurukshetra
Haryana, India

Abstract- Mining multiple level association rules in large databases is core area of data mining. Discovering these rules is favorable to the accurate and appropriate decision made by decision makers. Frequent patterns discovery is the key process in multiple levels association rule mining. One of the challenges in developing multiple level association rules mining approaches is to implement an iterative procedure to find association rule, which takes a intricate transaction process. Moreover, the offered mining methods cannot perform proficiently due to repetitive disk access overhead. Due to this, a novel method named MLTransTrie is presented in this paper. It can efficiently discover the association rules at multiple levels of abstraction in large databases and provide the more exact and precise information. The focus is on the comparative exploration and performance assessment of the MLTransTrie algorithm that generates multiple trie structure for all levels in one database scan. For this, the performance of this new algorithm MLTransTrie is analyzed and compared with LWFT and MLT2_L1 algorithm on diverse class of datasets (four real world dataset) and parameters.

Keywords- Concept Hierarchy; Confidence; Frequent Patterns; Minimum Support; MLTransTrie Implementation; Multiple-Level Association Rule.

I. INTRODUCTION

In a small span of time, remarkable rise has been observed, in terms of amount of data to be collected into databases, for a variety of computers applications, database tools and automated data gathering methods. The dramatic expansion in size of databases generates huge demand for analyzing such data and mining latent knowledge. Knowledge Discovery in Database (KDD) and Data mining appears as a solution to the problem of data analysis. Data mining is works as a step of KDD process. Data mining techniques are usually falls in to two categories: Descriptive and Predictive. The mean of the descriptive tasks are to extort formerly unknown and practical information which identifies patterns or relationships in data, while for the predictive ones, is to make prediction regarding data values using known results originated from different data. For the fulfillment of these objectives of data mining, a number of tasks include classification, prediction, regression, time series analysis, sequential pattern discovery and analysis, Clustering, Summarizations and Association Rule mining are used. In recent years, the expansion of data mining methods and approaches has received a great deal of consideration. It plays an important role in an extensive range of applications like market analysis, cloud computing, e-commerce, business, intrusion detection, manufacturing, etc.

Association rules mining is one of the most efficient technique of data mining. In large transactional databases, to find the association rules between items has been identified as a significant part of database research [1]. These rules can be efficiently applied to expose unknown associations, fabricating results that can make available a basis for decision making and forecasting. Basically the association rules mining deals with to find the relationship among buying patterns of diverse items of superstore data. At present, study on association rules is motivated by a variety of application areas, such as product manufacturing, business analysis, telecommunications and health care. It is also applicable in the area of creating statistical glossary from the text databases [2], sensor network data mining [3], discovering web access patterns [4], gene ontology mining [5], data mining in cloud computing [6], discovering associated images from large extent image databases [7], spatial data mining [8, 9] and also detection of intrusions from huge size network databases [10, 11].

In last few years, a variety of algorithms for mining association rule [12, 13, 14, 15, 16] has been designed which can be categorized into two approaches: first one candidate set generation approach as like Apriori [12] and other pattern growth approach [15,16]. Later than many scholars offered various enhancement of these approaches [17, 18, 19, 20, 21, 22, 23]. These offered methods were proposed to generate single level association rules which hold enormously broad information.

The existing methods have some negative aspects: Firstly, candidate set generation based mining algorithms are typically developed in form of some passes in order that the entire database needs to be scan from disks some times for each client's query. This is extremely inadequate in terms of reacting regarding the client's query rapidly. Secondly, an incredibly enormous number of association rules are generated by these existing algorithms. It is almost inconceivable for the end users to understand or certify such large number of association rules, by this mean restraining the effectiveness of the data mining results. Last but not least, to find appropriate amount of association rules the users have to apply an attempt and - error method because, no guiding information is given to opt suitable constraint parameters as confidence and support. Consequently, the process of mining association rules is converted into inefficient and exceptionally time intensive process. Therefore, it is also a challenge in designing proficient methods to handle large transactional databases.

The multiple level association rules came into existence to get more understandable and detailed facts. To facilitate finding of multiple level association rules, it is indispensable to offer data at multiple levels of abstraction. The data at multiple levels of abstraction can be provided with the help of concept hierarchy of items. These hierarchies describe the relationships between the items, and grouping data at several levels of abstraction. In various applications, the taxonomy information is either stored absolutely in the database or created by specialists in the application field or users. The arrangement of data at multiple levels of abstraction has become a common practice due to advancements in data warehousing and OLAP technology [24].

To explore effective methods of mining multiple-level association rules several promising directions are available. One approach is to use the uniform support and confidence thresholds for mining association rules at multiple levels. Other approach is to make use of non-uniform minimum support thresholds and probably different minimum confidence thresholds as well as mining associations at different levels of abstraction. In this approach value of minimum support threshold at primitive level is high and gradually decreasing at lower levels of abstraction. It helps to discover significant, informative association rules because of using non-uniform support thresholds at all levels. In this research work, non-uniform support and confidence thresholds have been used to avoid uninteresting rules, generated due to low threshold value.

The rest of the paper is divided in four sections. Section 2 presents the related work. Section 3 provides the brief overview of proposed algorithm. Results for running time for different datasets and comparisons with MLT2 and LWFT are discussed in section 4. Section 5 deals with conclusion and future scope.

II. STATE OF THE ART

A number of algorithms have been proposed for mining multiple-level association rules. Most of these algorithms are based on generation of candidate sets according to the minimum support threshold imposed, and then mining multiple-level association rules by discovering frequent itemsets within these levels.

Along with these algorithms, Show-Jane and Chen Arbee L.P. proposed the MLAPG (Multiple-Level Association Pattern Generation) algorithm to discover itemsets at all levels of abstraction [25].The basic idea of MLAPG is to construct an association graph by scanning of database once and after that traverses the graph to produce all frequent itemsets. Moreover, the size of database is gradually reduced by removing the items which are not frequent items at the earlier conceptual level. But at the last level, no further reduction of the database is required.

The AprioriNewMulti algorithm had been presented by N. Rajkumar, for mining multiple-level association rules where the minimum support for itemsets depends upon the length of itemset [26]. A new notion of multi minimum support was introduced in this algorithm. The value of minimum support is different for different lengths of the itemset. This algorithm is not based on concept hierarchy.

Cao H., [27] devised a vernal algorithm FAMML_FPT which is based on frequent pattern tree to extract the association rules at multiple levels. This method introduced the concept of repaired items and the cross-level repaired items, which is encouraging to generate FP-tree from lower levels to higher levels. In this algorithm, association rules discovering time is reduced due to no need of generating the candidate itemsets.

Pater Mirela and Popescu Daniela E. devised the ADA-AFOPT algorithm [28], to provide more accurate frequent itemsets by using the concept of non-uniform minimum support. First of all, this algorithm scans the original database to find frequent items and arranges them in increasing order of occurrence. After that, an AFOPT structure is created in second scan to signify the conditional databases of the frequent items. This AFOPT structure is used to compress the complete information regarding the transactional database. Later the itemset generation procedure is preceded by recursively traversing of the AFOPT structure.

Ramanaiah Preethi Kolluru, presented a new hybrid algorithm AC tree for mining multilevel association rules [29]. The algorithm AC tree combines Apriori and COFI tree to overcome the shortcoming of these conventional algorithms and make it appropriate for mining frequent itemsets at multiple levels. This method generally performs 2 key steps. During the initial step, Apriori algorithm is used to find out the frequent 1-itemsets by scanning the dataset. After that in next step, Fp header table is created by those 1- frequent itemset, subsequently these Fp trees for each element is considered for mining frequent itemset at each and every level.

Beside these approaches, Kousari Alireza Mirzaei Nejad et. al.[30] implemented an innovative multilevel fuzzy association rule mining approach for extraction of hidden knowledge. This approach uses the different minimum support threshold at all levels and different membership function for each item. In this method fuzzy boundaries are included instead of sharp boundary intervals and as a result of this, extracted rules, are more close to reality.

In order to improve the efficiency, Tang et al. [31] proposed an approach through grouping and merging the single level association rules generated by FP-growth. In this algorithm, hash table is used to enhance the efficiency of searching items. Conversely, when this approach is employed to analyze big data, the computation and memory overhead is increases exponentially which leads to a major bottleneck in big data analysis.

Furthermore Yang Xu presented an innovative genetic-based method to speed up the generation of multiple-level association rules [32]. In this algorithm, the tree encoding representation is used, which significantly decrease the association rule exploration space and also reduced the unnecessary computation.

III. DESCRIPTION OF MLTransTrie ALGORITHM

This new method is designed to fulfill the inevitability of a competent and effective algorithm to discover frequent itemsets at different levels of abstraction. The algorithm employs the framework of the advanced data structure Trie[33]. The idea of Trie is similar to frequent itemset mining algorithms such as FP-tree and COFI-tree. Both algorithms are similar in projection of each large item in the FP-tree data structure. The building of FP-tree needs two complete scan of database. The first scan produces the large1-itemsets and the second scan is used to arrange the all large items in descending order based on their support.

The proposed algorithm relies on transforming the whole database in transposed, reduced Boolean database. Therefore the space utilization of this algorithm is significantly reduced. Its key idea is based upon employing the scan reduction method, to locate new large itemsets with only single scan of database.

In practice, the proposed algorithm built a Trie with all transactions from transposed database, which is suitable for candidate set generation because transactions with same items use the same prefix tree. All the items in Trie are stored with their support. This algorithm discovers the set of all large itemsets by initially, finding the set of maximal itemset using a traversal technique and after that generates all subsets from this set with their support. So, this algorithm requires considerably small memory to generate large itemsets as compared to the existing FP-tree based methods.

The flowchart of the MLTransTrie algorithm is represented in Figure 1. The algorithm consists of two main procedures: the CREATE_TRIE procedure, which essentially applied for the generation of Trie with all items of transposed database and the GET_FREQUENT_ITEMSETS procedure, which is used to generate large itemsets form traversal of Trie. Initially, the proposed algorithm sets the level equal to maximum level and checks if the maximum level is equal to or greater than one. If the condition is true then it calls the CREAT_TRIE procedure of MLTransTrie algorithm for that level. Subsequently, the algorithm generates large itemsets by calling GET_FREQUENT_ITEMSETS procedure and decreases the value of max_level by one.

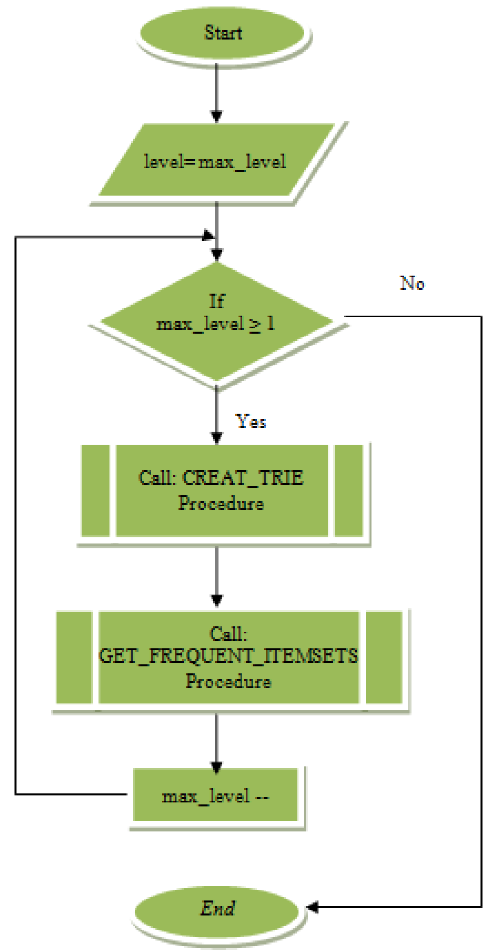

Figure 1. The flowchart of MLTransTrie Algorithm

Algorithm 1: MLTransTrie Algorithm

Inputs: DB, TD, CH and M_supp

Output: $FI_{(l,k)}$

1. Find the maximum level of CH.
2. For max_level to 1

For each level l

Call: CREATE_TRIE().

// Generation of Trie at each level.

For each level l

Call: GET_FREQUENT_ITEMSETS().

// Generation of frequent itemsets at each level.

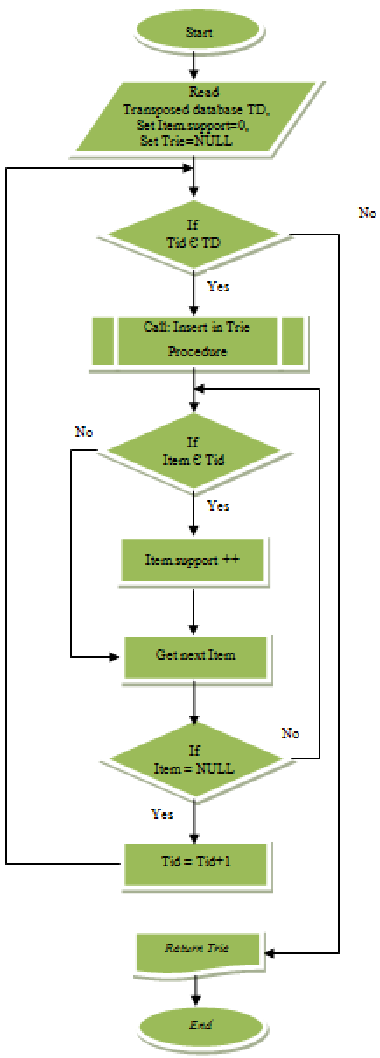

Figure 2. The flowchart of CREATE_TRIE Algorithm

Figure 2 illustrates the CREATE_TRIE procedure of the MLTransTrie algorithm. First, the algorithm read the transaction from transposed database and set the support value of item to zero. If the tid belongs to TD, call the insert in Trie procedure. On the other hand, for all items, if the item belongs to tid increase the support of item by one. After the completion of current tid get next tid from TD.

Algorithm 2: CREATE_TRIE procedure of MLTransTrie Algorithm

Input: TD, M_supp.

Output: T

//initially T= Ø, X.support= Ø.

1. For each transaction Tid

If Tid Є TD then

Insert Tid in T

//Adding all items of transaction in to Trie.

2. For each item X

If XЄTid then

X.support= X.support+1;

//Counting the frequency of item X.

3. return T;

// Return final Trie with frequency of all items.

Algorithm 3: GET_FREQUENT_ITEMSETS Procedure

Input: 1) T, M_supp.

Output: All frequent and infrequent itemsets.

1. GET_FREQUENT_ITEMSETS(Trie, M_supp) do
2. Frequent_itemsets=NULL; Suspected_itemsets= NULL:
3. For all (paths) routes in Trie do
4. Maximal_itemsets= DFS Traversing(Trie);
5. For all subset of Maximal_itemsets do
6. If itemset.support ≥ M_supp than do
7. Frequent_itemsets =Frequent_itemsetsU itemset;
8. End
9. Else than do Suspected_itemsets= Suspected_itemsetsU itemset;
10. End
11. For all itemset of Suspected_itemsets do
12. If itemset.support ≥ M_supp than do
13. Frequent_itemsets =Frequent_itemsetsU itemset;
14. Itmeset.support++;
15. End
16. End
17. End
18. return Frequent_itemsets;
19. return Suspected_itemsets;
20. End

Table 1 reveals the meaning of various symbols used in proposed algorithm MLTransTrie. This table is used to form better understanding of the algorithm.

TABLE 1. LIST OF SYMBOLS USED IN THE PROPOSED ALGORITHM.

Symbol	**Meaning**
DB	The original database
TD	The transposed Boolean database
CH	The concept hierarchy
M_supp	The minimum support
$FI_{(l,k)}$	The frequent k-itemsets at level l
T	The Trie
Tid	The transaction id
X	An itemset
X.support	The number of transactions containing X in the database

IV. EXPERIMENTAL STUDY

In this section, a performance comparison of the new algorithm MLTransTrie with some other multiple level algorithms is presented. In order to study the performance of new algorithm ML_TransTrie, implementation and testing is done on a Intel core i7 2 GHz machine with 6 GB of primary memory, running on Microsoft Window 8 and the program has been implemented in advance java using eclips.

In the accompanying subsections, first of all three real-world databases have been utilized to test the efficacy of the algorithm. After that, on the basis of parameters like support and confidence, it is compared with MLT2[34] and LWFT[35] and the outcome is demonstrated. Lastly, the analysis of the scalability of the method is introduced in this paper.

A. Databases

In order to get a real life result for proposed algorithm, it has been tested on three real world datasets. These real world datasets Breast-cancer, Credit-g and Soybean are available on UCI Repository of Machine Learning databases [36]. The brief description of datasets is given below.

1) *Breast-cancer:* It is a database concerning a study of the one of three domains provided by the Oncology Institute that has repeatedly appeared in the machine learning literature. Breast cancer domain was obtained from the University Medical Centre, Institute of Oncology, Ljubljana, Yugoslavia in years 1987-1988. This data set includes 201 instances of one class and 85 instances of another class. The instances are described by 9 attributes, some of which are linear and some are nominal.

2) Credit-g: Credit-g dataset is commonly used for result compilation of data mining algorithms. It shows information about German credit customers. There is 1000 number of transactions in dataset. This dataset contains 20 attributes in which 7 are numeric and 13 are categorical attributes. Several attributes that are ordered categorical have been coded as integer. Credit-g original dataset, is provided by Prof. Hofmann, contains categorical/symbolic attributes.

3) Soybean: Soybean is extensively accepted as a standard dataset for research work. It represents picture of soybean disease diagnosis data. This dataset contains 19 classes, only the first 15 of which have been used in prior work. The legend appears to be that the last four classes are unjustified by the data since they have so few examples. There are 35 categorical attributes, some nominal and some ordered. The value "dna" means does not apply. The values for attributes are encoded numerically, with the first value encoded as "0" the second as"1" and so forth. An unknown value is encoded as "?". The number of instances this dataset is 683. This dataset is donated by Ming Tan & Jeff Schlimmer on 11 July 1988.

B. Experimental set-up

For all the experiments conducted in this study, the transactional dataset is transformed into an encoded transactional database with the help of concept hierarchy. The concept hierarchies of datasets are provided by XML file. The minimum support thresholds are varies from 0.10 to 0.90. The support values are reduced as going from lower levels to upper levels, which are handled by delta factor. The parameter minimum confidence is taken a constant value for all results.

Figure 3, 4 and 5 illustrate the outcomes obtained by MLTransTrie on different datasets. After analyzing the results obtained through proposed algorithm it can state that running time decrease along with the increase of the minimum support. On the other hand, the running time also depends on number of items and number of transactions in database. The result shows that at constant minimum support running time is different for different dataset.

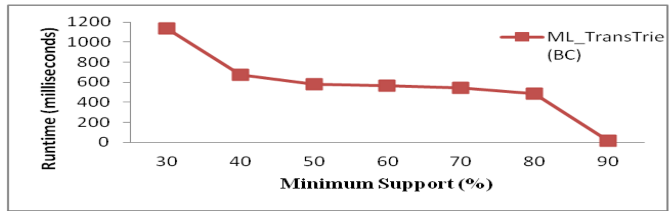

Figure 3. The run time on Breast-cancer dataset

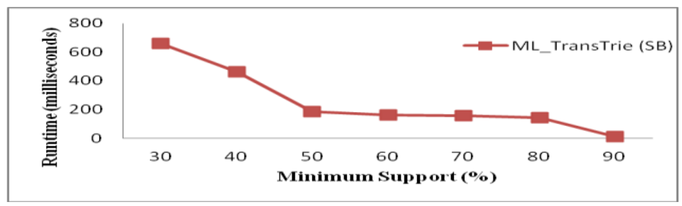

Figure 4. The run time on Soybean dataset

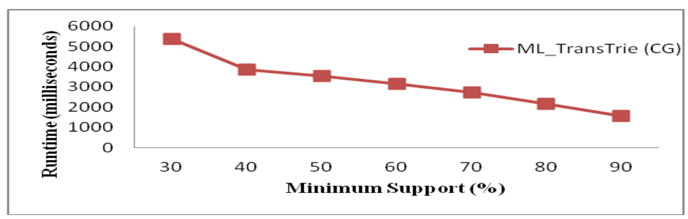

Figure 5. The run time on Credit-g dataset

The comparative analysis of MLTransTrie, MLT2 and LWFT have been presented in figure 6, 7 and 8, we can observe that the traditional algorithms expend the large amount of time in mining the frequent itemsets as compared to MLTransTrie. Moreover, in the situation of reduced minimum support the ratio of runtime between MLTransTrie and MLT2, LWFT is much less.

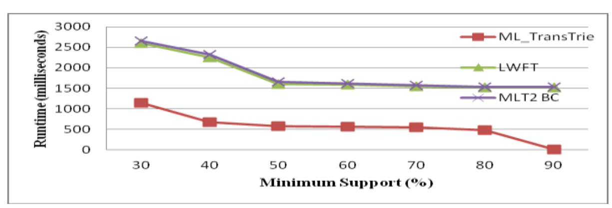

Figure 6. The run time on Breast-cancer dataset

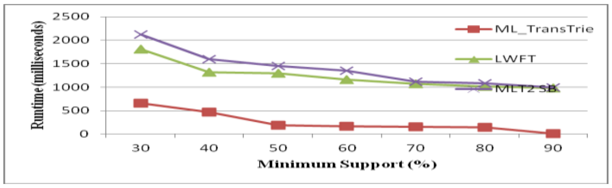

Figure 7. The run time on Soybean dataset

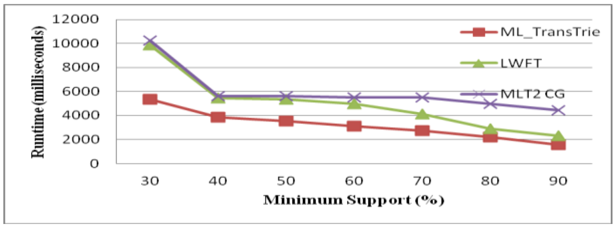

Figure 8. The run time on Credit-g dataset

Finally, it is worthwhile to state that the multiple level association rule extraction method MLtransTrie is better than existing algorithms in terms of time efficiency and space utilization. Nevertheless, the execution time of the multiple level association rule mining methods increase when the database size increases.

V. CONCLUSION

In this Paper, an innovative multiple-level association rule mining algorithm MLTransTrie is proposed which employed an advanced data structure Trie, multiple-level hierarchy, transposed database and non-uniform minimum support for each level to extract association rules at different levels of abstraction. The experimental results expose that: The proposed algorithm has the proficiency to find association rules at different levels according to the user's requirements, and this is due to using different minimum support threshold for each level. In addition, this method also reduce the space requirements because, it make use of transposed Boolean database.

This research work focuses on the comparative performance assessment of proposed algorithm and the algorithms presented in [34, 35]. The results expose that: In terms of discovering of association rules between the data items in large databases, the proposed method perform superior than previous methods. In particular, only one scan of the entire database is required during the execution of the algorithm, it is because of transformation of complete database in to Trie. It can accelerate the process of extracting considerably as demonstrated in the performance evaluation

However the rules taken out from proposed algorithm are desirable for users, but it is clear that with these attractive rules, some useless and redundant rules are also generated. Coping with this issue is a further attention-grabbing expansion of the existing work that requires some efficient optimized algorithms to reduce the unnecessary association rules.

REFERENCES

[1]. J. Han and M. Kamber, Data Mining: Concepts and Techniques, Morgan Kaufmann Publishers, 2000.

[2]. J. D. Holt and S. M. Chung, "Efficient Mining of Association Rules in Text Databases," CIKM'99, Kansas City, USA, pp. 234-242, Nov. 1999.

[3]. D. Han, Y. Shi, W. Wang, "Research on Multi-Level Association Rules Based on Geosciences Data", Journal of Software, vol. 8, no. 12, pp.3269–3276, 2013.

[4]. Mobasher, N. Jain, E.H. Han, and J. Srivastava, "Web Mining: Pattern Discovery from World Wide Web Transactions," Department of Computer Science, University of Minnesota, Technical Report TR96-050, March 1996.

[5]. P. H. Guzzi, M. Milano and M. Cannataro, "Mining Association Rules from Gene Ontology and Protein Networks: Promises and Challenges," In Proceeding 14th International Conference on Computational Science, Published by Elsevier Vol. 29, pp.1970-1980, 2014.

[6]. K.W. Lin and D.J. Deng, "A Novel Parallel Algorithm for Frequent Pattern Mining with Privacy Preserved in Cloud Computing Environments," International Journal of Ad Hoc and Ubiquitous Computing, Inderscience publication, pp.205-215, 2010.

[7]. R. Agrawal, T. Imielinski, and A. Swami, "Mining association rules between sets of items in large databases,"In Proceedings of the ACM SIGMOD International Conference on Management of Data (ACM SIGMOD '93), pp. 207-216, Washington, USA, May 1993.

[8]. Appice, M. Berardi, M. Ceci, and D. Malerba, "Mining and Filtering Multi-level Spatial Association Rules with ARES," Proceedings in 15th International Symposium, ISMIS 2005, Saratoga Springs, NY, USA, pp.342-353, 2005.

[9]. Petelin, I. Kononenko, V. Malaocioc, and M. Kukar, "Multi-level association rules and directed graphs for spatial data analysis," Expert Systems with Applications, vol. 40, no. 12, pp.4957–4970, 2013.

[10]. H. Zhengbing, L. Zhitang, and W. Jumgi W., "A Novel Intrusion Detection System (NIDS) Based on Signature Search of Data Mining," WKDD First International Workshop on Knowledge discovery and Data Ming, pp. 10-16, 2008.

[11]. H. Han, X. L. Lu, and L. Y. Ren, "Using Data Mining to Discover Signatures in Network-Based Intrusion Detection," Proceedings of the First International Conference on Machine Learning and Cybernetics, Beijing, vol. 1, 2002.

[12]. R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, and A. I. Verkamo. "Fast discovery of association rules: In Advances in Knowledge Discovery and Data Mining," pp.307-328. AAAI Press, 1996.

[13]. R. Bayardo and R. Agrawal. "Mining the most interesting rules," In Proceedings of the 5th International Conference on Knowledge Discovery and Data Mining (KDD '99), pp.145-154, San Diego, California, USA, August 1999.

[14]. J. Hipp, U. Güntzer, and U. Grimmer. "Integrating association rule mining algorithms with relational database systems," In Proceedings of the 3rd International Conference on Enterprise Information Systems (ICEIS 2001), pp. 130-137, Setúbal, Portugal, July 2001.

[15]. R. Ng, L. S. Lakshmanan, J. Han, and T. Mah. "Exploratory mining via constrained frequent set queries," In Proceedings of the 1999 ACM-SIGMOD International Conference on Management of Data (SIGMOD '99), pp. 556-558, Philadelphia, PA, USA, June 1999.

[16]. Y. Guizhen, "The complexity of mining maximal frequent itemsets and maximal frequent patterns," Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining , pp. 343-353, Seattle, WA, USA, August 2004.

[17]. J.S. Park, M.S. Chen and P.S. Yu, "An Effective Hash Based Algorithm for Mining Association Rules," In Proceeding 1995 ACM SIGMOD International Conference on Management of Data, pp.175-186, 1995.

[18]. Cheung, J. Han, V.T. Ng, A.W. Fu and Y. Fu, "A Fast Distributed Algorithm for Mining Association Rules," In Proceeding of 1996

International Conference on Parallel and Distributed Information Systems (PDIS'96, Miami Beach, Florida, USA), 1996.

[19]. S. Brin, R. Motwani, J.D. Ullman, and S. Tsur, "Dynamic Itemset Counting and Implication Rules for Market Basket Data," In Proceedings of the 1997 ACM SIGMOD International Conference on Management of Data, Vol.26 No. 2, pp.255–264, 1997.

[20]. M.J. Zaki, "Scalable Algorithms for Association Mining," IEEE Transactions of Knowledge and Data Engineering, Vol. 12, pp.372–390, 2000.

[21]. B. Vo, H. Nguyen, T. B. Ho and B. Le, "Parallel Method for Mining High Utility Itemsets from Vertically Partitioned Distributed Databases," Proceedings of the 13th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, pp.28-30, 2009.

[22]. K. Shrivastava, P. kumar and K. R. pardasani, "FP-Tree and COFI Based Approach for Mining of Multiple Level Association Rules in Large database," International Journal of Computer Science and Information Security (IJCSIS), Vol.7, No. 2, 2010.

[23]. M. Rajalakshmi, T. Purusothaman and R. Nedunchezhian, "Maximal frequent itemset generation using segmentation approach," International Journal of Database Management Systems". Proceeding in IJDMS, Vol. 3, No. 3, pp.19-32, 2011.

[24]. S. Chaudhuri and U. Dayal, "An Overview of Data Warehousing and OLAP Technology," ACM SIGMOD Record, vol. 26, pp. 65-74, 1997.

[25]. Show-Jane and Chen Arbee L.P., "A Graph-Based Approach for Discovering Various Types of Association Rules," Proceeding in IEEE Transactions on Knowledge and Data Engineering. Vol. 13 No. 5, pp 839-845, 2001.

[26]. N. Rajkumar, M.R. Karthik and S.N. Sivanandam, "Fast Algorithm for Mining Multilevel Association Rules," Proceeding in IEEE Transactions on Knowledge and Data Engineering, vol. 02, 2003.

[27]. H. Cao, Z. Jiang, and Z. Sun, "Fast Mining Algorithm for Multilevel Association Rules Based on FP-tree," Computer Engineering, vol. 19, no. 25, 2007.

[28]. Pater Mirela and E. Popescu Daniela, "Multi-Level Database Mining Using AFOPT Data Structure and Adaptive Support Constrains," Proceeding in Int. J. of Computers, Communications & control, ISSN 1841-9836, Vol. 3, pp 437-441, 2008.

[29]. Ramanaiah Preethi Kolluru, "Hybrid Association Rule Mining Using AC Tree". Proceeding in Journal of Information Engineering and Applications, Vol. 1, No. 02, 2011.

[30]. Kousari Alireza Mirzaei Nejad, Mirabedini Seyed Javad, Ehsan Ghasemkhani, "Improvement of Mining Fuzzy Multiple-Level Association Rules from Quantitative Data," In Proceedings Journal of Software Engineering and Applications Vol. 5, No. 3, pp 190-199, 2012.

[31]. H. Tang, M. Wu, and Y. He, "Improved multilevel Association Rule Mining Algorithm," Computer Engineering, vol. 16, No. 16, 2011.

[32]. Y. Xu, M. Zeng, Q. Liu and X. Wang, "A Genetic Algorithm Based Multilevel Association Rules Mining for Big Datasets," Hindawi Publishing Corporation Mathematical Problems in Engineering, Vol.9, 2014.

[33]. F. Bodon and L. Ronyal, "Trie: An Alternative Data Structure for Data Mining Algorithms," Proceeding in Mathematical and Computer Modeling, Elsevier Ltd., Vol. 38, pp.739-751, 2003.

[34]. J. Han and Y. Fu, "Discovery of Multiple-Level Association Rules from Large Databases," In Proceedings of the 21st International Conference: Very Large Data Bases, vol. 95, pp.420–431, 1995.

[35]. M. Vishav, R. Yadav and D. Sirohi, "Mining Frequent Patterns with Counting Inference at Multiple Levels," Proceeding in International Journal of Computer Applications Vol. 3, No.10, PP. 1-6, 2010.

[36]. http://repository.seasr.org/Datasets/UCI/arff/