Volume No:4, Issue No:10 (March-2019)
ISSN No : 2454-423X (Online)
International Journal of Research in Advanced Computer Science Engineering
A Peer Reviewed Open Access International Journal
www.ijracse.com

An Adaptive Data Mining of Frequent Item Sets on Sublunary Data Using FP Miner

P.Nithyasri
Department of Computer Science & Engineering,
NS Raju Institute of Technology,
Visakhapatnam, Andhra Pradesh-531173, India.

K.Suryatej
Department of Computer Science & Engineering,
NS Raju Institute of Technology,
Visakhapatnam, Andhra Pradesh-531173, India.

Ch. Chetanaditya
Department of Computer Science & Engineering,
NS Raju Institute of Technology,
Visakhapatnam, Andhra Pradesh-531173, India.

G. Rajasekharam
Department of Computer Science & Engineering,
NS Raju Institute of Technology,
Visakhapatnam, Andhra Pradesh-531173, India.

Abstract
Data mining is the procedure for finding out correct patterns and knowledge from large data sets and it is also used for temporal data mining. The categories are as prediction, classification, clustering, search& retrieval and pattern discovery. The work of those categories is considered with temporal data Temporal data can hold time-stamped information that affects the results of data mining. The main focus lies in the generation of frequent patterns using association rule mining algorithms like Apriori, Dynamic Item set Counting, FP-growth, and Matrix Based Association Rule Mining Algorithm.

Keywords
Temporal data, Data mining and Frequent item sets.

1. Introduction
Data mining is the procedure for finding out correct patterns and knowledge from large data sets. The main aim of data mining process is to mine knowledge from current dataset and change it into the human understandable form to advance use. There are various data analysis techniques available for research studies. These analysis techniques include temporal data mining, association rule mining, frequent item set mining. The motivation behind this paper is to explore data mining techniques, which is suitable for finding frequent item sets using temporal data.

It also uses synthetic datasets for showing a quite effective experiment result of proposed algorithm. It discovers frequent item sets with respect to an optimal time period. The aim of this paper is to propose and design system for frequent item set mining on temporal data. The proposed solution uses an algorithm of frequent item set mining with time cube. This work proposes an efficient flow to find out interested or frequently utilized item sets from the database. But discovered patterns may not be valid [1], because there is no any proof which will show that a particular pattern hold at particular time interval. This will arises the problem of Over estimation [1]. Association Rule mining is one of the fundamental data mining method. Agarwal first introduced the problem of association rule mining in 1993. Since then it is one of the most popular research area on the field of knowledge discovery. The association rule-mining problem is commonly known as the market basket analysis, but there are several applications that use association rules as well i.e. biological research areas, telecommunication and network analysis etc. Regarding the diversity of the applications that use association rule mining, several algorithms have been developed.

Cite this article as: P.Nithyasri, K.Suryatej, Ch. Chetanaditya & G. Rajasekharam, "An Adaptive Data Mining of Frequent Item Sets on Sublunary Data Using FP Mine", International Journal of Research in Advanced Computer Science Engineering, Volume 4, Issue 10, 2018, Page 1-5.

Volume No:4, Issue No:10 (March-2019) ISSN No : 2454-423X (Online)
International Journal of Research in Advanced Computer Science Engineering
A Peer Reviewed Open Access International Journal
www.ijracse.com

All of these algorithms have their own advantages and disadvantages, so it is useful to compare them. Association is a rule, which implies certain association relationships among set of objects such as occur together or one implies the other. Goal of Association rule mining helps in finding interesting association relationships among large set of data items. The discovery of such associations can help develop strategies to predict. An Association Rule is a rule of the format, LHS (left hand side) => RHS (ride hand side). Let X and Y are two item sets where X, Y $\subseteq$ I (both side contains sets of items) and X$\cap$Y =Ø (don't share common items). Briefly, an Association Rule is an expression; X $\Rightarrow$ Y, where X and Y are set of items.

Each rule is assigned two factors: 1) Support and 2) Confidence. So to avoid this problem new threshold is used called as density. Also, it refers recommendation for item sets having high occurrence value or support. A New Methodology for Mining Frequent Itemsets on Temporal Data [1], the main aim is to develop an efficient algorithm by expending a priori algorithm. The idea of time cubes is used to handle time hierarchies. The patterns will consider the way by which the patterns that happen periodically, during a time interval or both, are recognized [1]. A new value of threshold as a density is also proposed to solve the problem of overestimating with respect to time periods and also make sure that discovered patterns are valid [1].

AprioriMin [2], which presents a new algorithm named AprioriMin Algorithm which mines the frequent itemsets [2]. This algorithm aims at introducing a new strategy for the calculation of frequent itemsets to prune phase of frequent itemsets. This technique decreases the execution time, when the support threshold increases for AprioriMin Algorithm [2]. **N**ew model which integrates topic modeling and enhanced frequent itemsets mining together to alleviate the excellent strength on enhancing the information filtering as well as the reducing time consuming of the overall process [3]. Proposed technique is given for mining fuzzy association rules that have a temporal pattern. The result of search was considered by examining the temporal association rules and composition of fuzzy, it also combined a multi objective evolutionary algorithm with iterative rule learning to mine many rules [4]. Mining fuzzy periodic association rules they develop techniques for discovering patterns with periodicity in this work. Patterns with periodicity are those that occur at regular time intervals, and therefore there are two aspects to the problem: finding the pattern, and determining the periodicity [5]. The difficulty of the task lies in the problem of discovering these regular time intervals, i.e., the periodicity [5].

It relies on the extraction of frequent item sets from usage databases. However, those data bases are usually considered as a whole, and therefore, item sets are extracted over the entire set of records. Our claim is that possible subsets, hidden within the structure of the data and containing relevant item sets, may exist [6]. Discovery of Fuzzy Temporal Association Rules they propose a data mining system for discovering interesting temporal patterns from large databases. The mined patterns are expressed in fuzzy temporal association rules which satisfy the temporal requirements specified by the user.

2. Implementation modules

2.1 Existing System:

In the existing system three different algorithms apriori, FP-growth, eclat for frequent item set mining are given. They have developed their own best combination method for finding frequent item sets from large database.

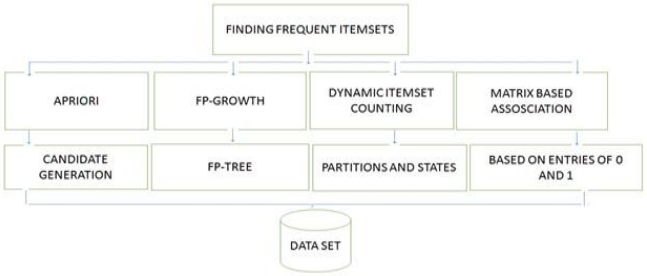
2.2 Proposed System:

The FP-miner is mainly aimed at identifying the frequent patterns from the mushroom dataset using different algorithms along with a newly proposed

Volume No:4, Issue No:10 (March-2019) ISSN No : 2454-423X (Online)
International Journal of Research in Advanced Computer Science Engineering
A Peer Reviewed Open Access International Journal
www.ijracse.com

algorithm. The execution times of these algorithms are compared and the best of them is identified.

2.3 Architecture:



2.4 Requirements:

Hardware Requirements:

- 512mb Ram
- Pentium 4
- 80gb Hdd

Software Requirements:

- Operating System – Windows
- Technology – Java
- Web Technologies – Html,Javascript,Css
- Database - Mysql

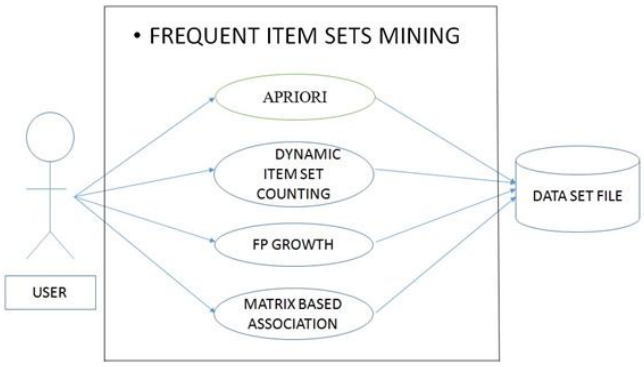


Fig. 1 Usecase Diagram

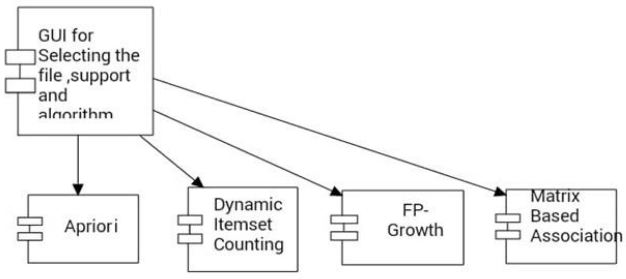


Fig. 2 Component Diagram

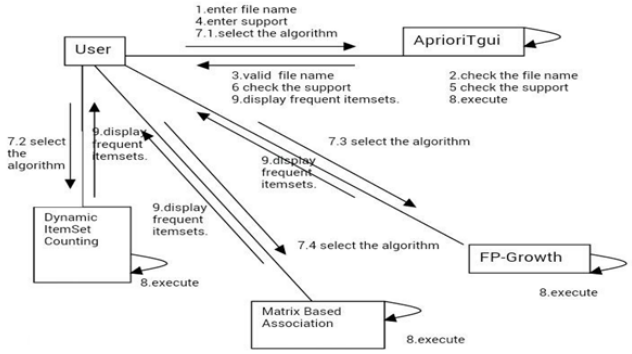
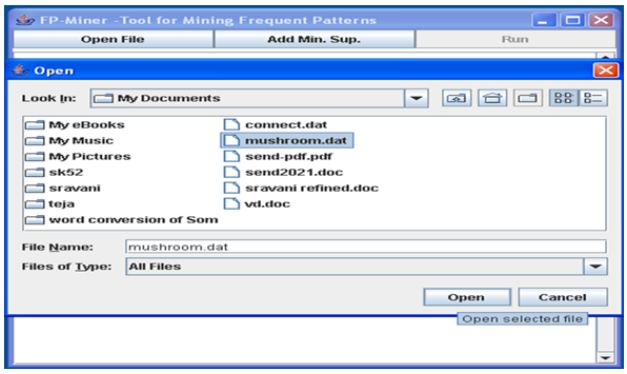


Fig. 3 Collaboration Diagrams:

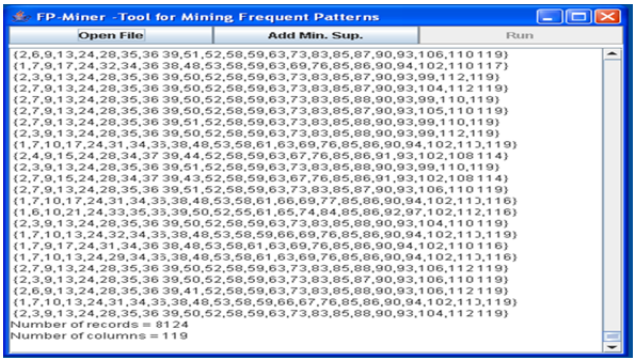
3. Result and Discussions

3.1 Coding-Output:

3.1.1 Step 1: Selection of File Name

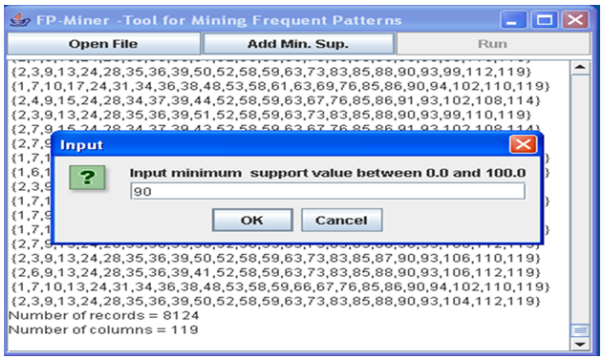


3.1.2 Step 2: Display the Contents Of The File On The Text Area

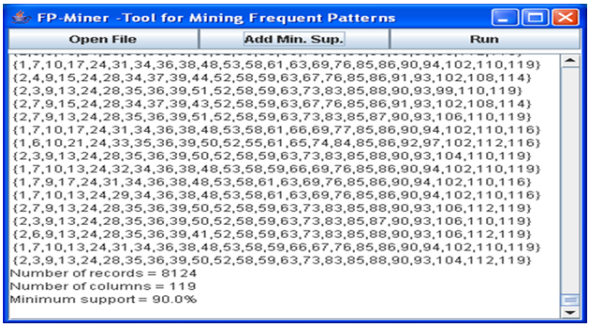


Volume No:4, Issue No:10 (March-2019) ISSN No : 2454-423X (Online)
International Journal of Research in Advanced Computer Science Engineering
A Peer Reviewed Open Access International Journal
www.ijracse.com

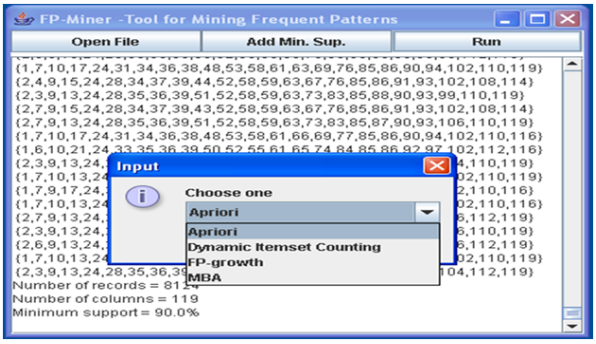
3.1.3 Step 3: Enter valid support



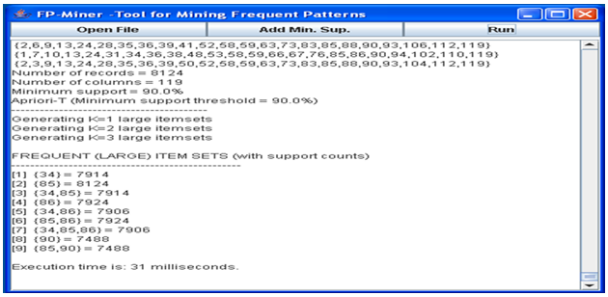
3.1.4 Step 4: Display valid support



3.1.5 Step 5: Select the algorithm



3.1.6 Step 6: Display the frequent patterns for apriori



3.2 TEST CASES :

Testcase Description	Support Value	Test Case Steps	Expected Execution Time of Apriori	Actual Execution Time of Apriori	Expected Execution Time of Fp-Growth	Execution Time Of Fp-Growth	Pass/ Fail
Enter valid support value	50	Select the algorithm	187ms	189ms	196ms	197ms	Pass
Enter valid support value	90	Select the algorithm	110ms	113ms	176ms	179ms	Pass
Enter valid support value	70	Select the algorithm	78ms	77ms	148ms	144ms	Fail
Enter valid support value	110	Cant select algorithm	Nil	Nil	Nil	Nil	Fail

4 Conclusion:

Frequent Pattern mining is used for finding frequent itemsets among items in a given data set. An objective of frequent pattern mining is to develop a systematic method using the given data set and find frequent items. This project work is focused on explaining the fundamentals of association rule mining and analyzing the implementation of the well known association rule algorithms by comparing the execution time for generating frequent item sets with the different minimum support values. Study focuses on algorithms Apriori, FP-Growth and Dynamic Itemset Counting. At the same time, we described a new approach for association rule mining based on matrix based association rule mining. The results show that Apriori cannot be run very effective than FP Tree. Apriori on the other hand runs too slow because each transaction contains density data. DIC (Dynamic Itemset Counting) is much slower than every other algorithm for the real -dataset.

5 References:

[1] Data Mining and Concepts by Jiawei Han, Micheline Kamber, Morgan Kauffman Publishers.

[2] Data Mining Techniques by Arun K Pujari

[3] Object oriented Software Development using UML and Java by Timothy C Lethbridge and Langaniere, Mc grawhill Co.,

[4] Object Oriented Systems Development by Ali Bahrami

[5] Software Engineering –a practioner's Approach by Roger S Pressman, Mc grawhill Pn.

[6] The Complete Reference Java 2 by Herbert Schildt, Mc Grawhill Pn

[7] "An Analysis Of Association Rule Mining Algorithms" by Renate Iváncsy, Ferenc Kovács, István Vajk

[8] "Mining Association Rule Algorithms In Large Databases" By Semih Utku

[9] "Fast Algorithms for Mining Association Rules" Rakesh Agrawal , Ramakrishnan Srikant Proceedings of the 20th VLDB Conference Santiago, Chile, 1994

[10] Real World Performance of Association Rule Algorithms,Zijian Zheng Ron Kohavi, Llew Mason.