A Hybrid Approach for Re-current Frequent ItemSet Mining
Surendra Tiwari1 Prof. Suma B.2
1P.G. Student 2 Assistant Professor
1Department of Computer Science & Engineering
2RVCE, Bangalore, India.

Abstract— Recurrent itemsets mining is an important concept in the mining of different hidden patterns and is in demand in many real-life applications. There are many methods to find recurrent itemset based on breath first search (for ex-Apriori) and depth first search (for ex-FP-Tree). Both these have their own disadvantage and advantage. The BFS gave better result in terms of efficiency. The DFS algorithm is good in terms of time taken to generate recurrent itemset. Most of previous and current works on mining recurrent itemsets can only capture the static itemsets they do not take care of the recurrent itemsets that often change.so this limitation motivate to work in this area.In the proposed work, using the concept of pheomonnes Strategy of Ants, we described and proposed a new method for dynamic recurrent itemsets mining.

Keywords: Hybrid Approach, Apriori Algorithm, FP-Tree,

I. INTRODUCTION
Recurrence itemsets mining is the important thing to know many feature of data stored.in data mining individual tried to find out the hidden relationship between data stored. now a days most of organization are interested to know about hidden patterns ,so individual can conclude that most of global organization have a team related to data mining, so many approaches up to now are proposed for finding hidden patterns. with the developing and more detailed of the research on recurrent itemsets mining, it is widely used in the field of data mining, for example, mining association rule, correlation analysis, classification, and clustering et al. in the past scenario of data mining approaches, there are two types of mining methods were considered and taken through: Breadth-First Search algorithms and Depth-First Search algorithms. both these have their own disadvantage and advantage. the BFS gave better result in terms of efficiency. The DFS algorithm are good in terms of time taken to generate recurrent itemset. Examples of BFS and DFS are apriori(BFS) and FP-Growth(DFS).but these both algorithm are inappropriate for generating dynamic itemset. so generating dynamic itemset . we started to working with ant colony with appriori algorithm. But again this algorithm takes much more time to generate recurrent itemset.so outcome of all this to find a new hybrid approach that can take dynamic nature as well as time as a constrain in efficient,in this paper we presented a approach to find recurrent itemset using fp growth with ant colony algorithm.at the end experimental result showed that it is better than Apriori with ant colony approach.

The structure of the paper is as follows: Section 2 describe problem, section 3 describe hybrid approach for mining dynamical recurrent itemsets and results, and section 4 related to conclusion. Section 6 is related to references.

II. PROBLEM STUDY

Given a transaction dataset \{T\},and there are n transaction \{t1,t2,t3,...,tn\}.now every transaction consist of itemset which are determine by itemset. The item in transaction are from itemset\{i1,i2,i3,...,in\}.now for knowing the customer behavior we have to know the itemset of transaction which are frequently occurring.so there are many algorithm are proposed to solve this problem in next paragraph all previous algorithm detailed are there.

III. PROPOSED ALGORITHM AND RESULT
These following steps we followed during the work to get the things done:

1) Read Transaction File and convert the transaction file into matrix form
2) Now read the number the matrix row by row and calculate items. And determine the itemset and number of total items in file.
3) Now using Min_Support reduce the number of item and then sort that list using their support in decreasing order.
4) Now using this list sort all transaction and start from first item and if it is not in tree add it and again using FP tree process build Fp-tree and on node update two thing pheromone value and count.
5) Calculate the dynamic recurrent itemset using FP-tree and pheromone value which is there on every node.

Results what we get is 95% accurate. Here we are using supermarket transaction file.The dataset considered for the proposed approach is retail transaction dataset. Every row of retail transaction dataset is called a transaction that contains multiple columns which are called items of transaction. The number of item in transaction dataset may be varying from transaction to transaction. Results what we get is 95% accurate. In below table 4.1, comparisons between Eclat, FP-Growth and proposed approach are given.

<table>
<thead>
<tr>
<th>Miner</th>
<th>Proposed</th>
<th>Eclat</th>
<th>FP-Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>95.12</td>
<td>92.68</td>
<td>85.37</td>
</tr>
<tr>
<td>Precision</td>
<td>92.31</td>
<td>87.50</td>
<td>84.62</td>
</tr>
<tr>
<td>Recall</td>
<td>100.00</td>
<td>100.00</td>
<td>91.67</td>
</tr>
<tr>
<td>Specificity</td>
<td>88.24</td>
<td>100.00</td>
<td>76.47</td>
</tr>
</tbody>
</table>

Table 1: Comparisons between Eclat, FP-Growth, Proposed Approach

In the table 4.1 ,it is clear that proposed approach is more accurate then standard frequent itemset mining algorithms.
II. CONCLUSION
This work is done to analyze the customer behavior to increase the supermarket revenue. This is an hybrid approach that taken the advantage of both algorithm FP-tree as well as Ant Colony. In the paper we provided a perfect evaluation function with minimum overhead. The proposed algorithm does require many arithmetic operations but that make the algorithm to be proved numerically. A distribution of itemsets over all tree node is inexpensive. We defined a new data structure to implement a frequent itemset structure. The implementation is aimed to reduce the size of normal implementation of any technique that has been used. Also, using this data structure may recreate the tree by deleting all nodes of non frequent itemsets after a scanning a specific percentage of database.

REFERENCES