Finding Utility Pattern for Building Effective Ration Shop

S. Aswini¹, U. Deepikaa², M. Meena³, P. Krishnamoorthy⁴
Student¹,²,³, Assistant Professor⁴
Department of Computer Science and Engineering
Kingston Engineering College, Vellore, India

Abstract:
The process of mining high utility itemsets from transactional databases which is also a business intelligence technique refers to finding the itemsets which creates high demand to owners. Here in our case we are implementing this process government ration shop management process to identify and extract the food item requirement and purchase pattern to find the demand of items in different time intervals. All purchase of food items are registered in the centralized database which is going to be our transactional data for finding the high utility item set. SMS notifications are sent to the concern citizen when a transaction is made on their ration card. Also they will receive notification when the food items reach the warehouse along with their schedule to purchase them. We extract the transaction region wise and process this transaction to identify the pattern of the particular ration shop. We use FP growth algorithm tree based algorithm to find the high utility item pattern which will be used by the authorities to make decisions of sending the quantity food stock to ration shops.

Index Terms: High utility item set, frequent pattern, up tree.

I. INTRODUCTION

Government Ration shops are established to serve Indian citizen by providing the food materials for affordable cost. But due to delay in supply all citizen need to come to the ration shop and ask them whether they are providing the items today. Because of this problem when the items are provided people stand in big queue fighting for the items thinking that they won’t get items. If they are aware of the items content and when they are provided then it will be helpful for them. Our system not only intimated citizens about their sales but also provides schedule for buying the food items. When the food items are received in the ration warehouse the schedule will be sent as an SMS. Our system is centralized so all the data are stored in the central server when food items received and also at the time of sale. They data are processed using frequent utility item set mining algorithm to find the item demand pattern. These patterns are used to dynamically adjust the supplies quantity without wasting the food items and also not making them unavailable.

II. ARCHITECTURE

![Architecture Diagram]

Figure 1. Architecture
Extract Transaction Log
1. This module provides an interface to individual ration shop authorities to make a bill for each user transaction
2. These billing details from each ration shop are collected and stores in a central data server.
3. For ease of processing each items in the bill are assigned with a label and each bill is converted into a pattern

Utility Tree Construction
1. For constructing a utility tree we first extract utility (count of sale) value of each food item
2. After finding the utility value items are sorted in descending order to form an utility index table holding item label and value
3. With use of transaction table and the utility index table we construct the utility tree by inserting patterns one by one
4. utility value is assigned as weights for each item in the tree

Extract High Utility itemset
1. This module is responsible for extracting the final utility pattern from the utility tree
2. it uses the UP growth algorithm for finding the utility patterns by traveling downward in the utility tree
3. It travels downward toward the item with higher weight till the end of the utility tree
4. When the travel process reaches the end of the tree the pattern will be extracted completely

Scheduling Stock
1. After extracting the high utility item patterns for individual zone the required stock is calculated
2. Based on the calculated stock count the items are delivered to the individual ration shops
3. Because we have a billing process centralized the previous stock details is also take from the data base for calculating the demand

Scheduling Purchase
1. We have complete user details in the repository based on that information user are intimate their schedule of purchase
2. Based on the schedule user will purchase their food items and these details is also stored in the database for further analysis
3. This process avoid the big queue and reduce the common people work load and irritations

III. ALGORITHM
- UP Tree
- FP Growth algorithm
- Association Rule

UTILITY PATTERN TREE (UP Tree)
This algorithm for mining high utility itemsets with a set of techniques for pruning candidate itemsets. It efficiently maintains the information of transaction database related to the utility pattern. It generates the high utility item sets depending on construction of a global UP-Tree.

Algorithm UP Tree
Input: transactions T, items Ia
Output: UPTee
Initialization tree TR, DemandIndex PI
Foreach transaction ti in T
    Foreach item I in Ia
        itemDemand Di = SaleCount(ti)
        InsertIndex(Pi, Di)
    End Foreach
Sort(Pi)
Foreach transaction ti in T
    Foreach index x in DemandIndex Pi
        If (TR contains ti)
            AssignWeights(ti)
            Insert( TR, ti)
        Else
            AssignWeights(ti)
            Update( TR, ti)
        End
    End Foreach
End Foreach
Return TR

FP GROWTH ALGORITHM
- The FP Growth Algorithm is a way to find frequent itemsets without using candidate generations, thus improving performance. It retains the itemset association information. It does not generate the candidates, it only tests.
- If the frequent items can be stored in a compact structure then the original transaction database does not need to be used repeatedly.
- It scans the transaction database once, to find all frequent items.

Step 1: Build a compact data structure called the FP-tree. Built using 2 passes over the data-set.
Step 2: Extracts frequent itemsets directly from the FP-tree.

Advantages of FP-Growth
- only 2 passes over data-set
- It compresses the data-set
- no candidate generation
- much faster than Apriori

IV. LITERATURE SURVEY
- C. H. Cai, A. W. C. Fu, C. H. Cheng, and W. W. Kwong, “Mining association rules with weighted items,” in Proc. Int. Database Eng. Appl. Symp., 1998, pp. 68–77 . In this paper Association rule are often written as x>y meaning that whenever x appears y also tends to appears. x is often referred to as the rule’s antecedent and y as the consequent.
In this paper the pruning process is proposed to eliminate the item set which are not found to be frequent. This reduces the size of the tree and to process efficiently.

- R. Agrawal, R. Srikanth, “Frequent item set using a prior algorithm”. In proc. 20th Int. Conf. Very Large Databases, 1994, pp. 487–499. In this paper the algorithm for finding association rules was proposed. It is processed in two parts. Generating candidate item and pruning process.

V. CONCLUSION

This paper presents a survey on various High Utility Itemsets algorithms that were proposed by earlier researches for the better development in the field of Data Mining. Various algorithms and methods discussed above will help in developing efficient and effective High utility itemsets for data mining. Our main objective is to build an effective system to manage and deliver food items to Indian citizen by processing their past transactions. Our project can enhance Indian ration shop stock maintenance and delivery scheduling procedures. It also avoids delay in delivering and wastage of stocks. SMS notification reduces misuse of the food stock.

VI. REFERENCES


