A Weighted Frequent Itemset MiningAlgorithm for Intelligent Decisionin Smart Systems

Abstract

Data mining technology has been playing an increasinglyimportant role in decision making activities. Such as Frequent itemset mining (FIM), as an

important step of association rule analysis is becoming one of the most important research fields in datamining.FIM widely used in the field of precision marketing, personalized recommendation, network optimization, medical diagnosis and so on.

Weighted FIM in uncertain databases should take both existential probability and importance ofitems into account in order to find frequent itemsets of great importance to users.The weighted frequent itemsets not satisfy the downward closure property any longer.The search space of frequent itemsets cannot be narrowed according to downward closure property which leads to a poor time efficiency.TheWeight judgment downward closure property-based FIM (WD-FIM)algorithm is proposed to narrow the searching space of the weighted frequent itemsets and improve the time efficiency.The evolution of segmentation was supported by advancements in technology. The shift into digital enabled an easier capture and retention of data while increasingly efficient databases facilitated the usability of that data. Although advancements in technology were crucial to the type of market segmentation used in precision marketing, they were not the driving force behind it. Instead, customer demand and expectation, alongside the fierce competition, were the driving factors.

Existing System

To mine frequent itemsets from uncertain datais to apply the candidate generate-and-test paradigm.State-of-the-art algorithms based on tree structures cancause fatal problems in terms of runtime and memory usageaccording to the characteristics of uncertain databases andthreshold settings because their own tree data structures canbecome excessively large and complicated in their miningprocesses. Various approaches have been suggested to over come such problems.

Proposed U-Apriori algorithm whichapplies the candidate generate-and-test process to mine frequent itemsets from for uncertain data. Similar to Apriorialgorithm for mining precise data, U-Apriori algorithm needsto scan the database frequently and generates a large numberof candidate frequent itemsets.

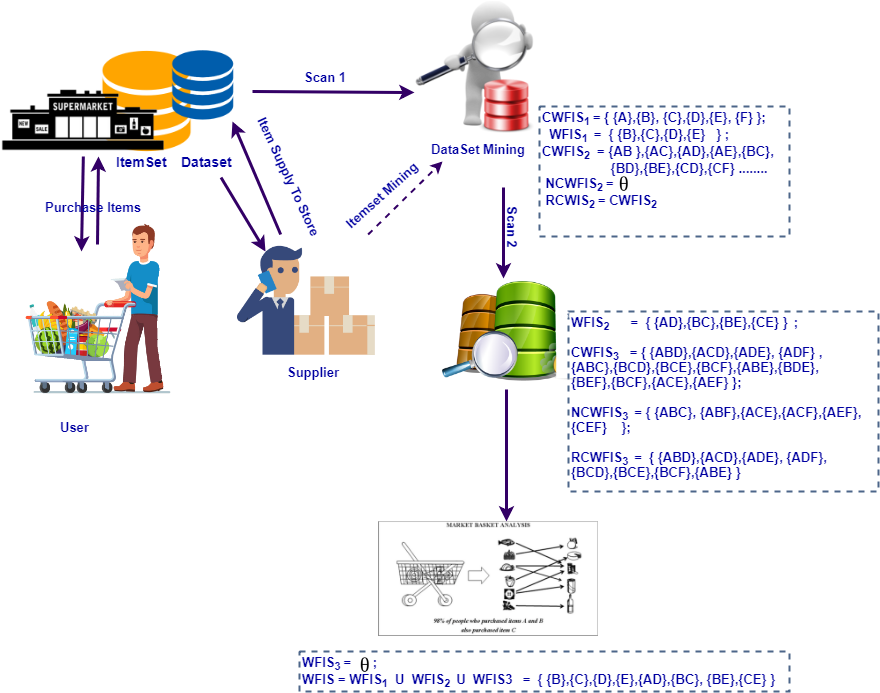
Proposed System

The proposed WD-FIM algorithm mines the weighted frequent itemsets from an uncertain database using the candidate generate-and-test paradigm. The weighted frequentitemsets are discovered by repeated iteration like U-Apriorialgorithm. Obviously, there are significant differencesbetween WD-FIM algorithm and U-Apriori algorithm. First,WD-FIM algorithm is proposed for mining weighted frequentitemsets in uncertain datasets. However, U-Apriori can onlybe used to discover frequent itemsets in uncertain datasets.Second, the basis of the proposed WD-FIM algorithm is theaforementioned weight judgment downward closure propertyand existence property of weighted frequent subsets, butthe downward closure property is used directly to narrowthe searching space of frequent itemsets in U-Apriori algorithm. Based on the aforementioned definitions and theorems,the pseudo code of proposed WD-FIM algorithm.

Future Work

In Future work The WD-FIM algorithm used resultstakes the input as: an uncertain transactional dataset, *DS*;a weight table, *wtable*; a user-specified minimum expectedweighted support threshold,Then itscans the dataset to get the weighted frequent 1-itemsets. Finally, on the basis of the proposed weightjudgment downward closure property and the existenceproperty of weighted frequent subsets, the weighted frequent*k*-itemsets will be discovered and all weighted frequentitemsets will be returned. During thisprocess, the calculation of *CWFISk* is extremely is performed to guarantee all theweighted frequent *k*-itemsets are included in *CWFISk* according to the existence property of weighted frequent subsets.

Architecture



Modules

User

Users Buying goods and the services from merchants who sell on the Internet. Since the emergence of the World Wide Web, Shoppers can visit web stores from the comfort of their homes and shop as they sit in front of the computer Consumers buy a variety of items from online stores. In fact, people can purchase just about anything from companies that provide their products online.

Store Market

Merchants have sought to sell their products to people who surf the Internet.

Before people buy anything online, get to know the seller people needto know their contact detailsfora reputablebusiness should make this information easy to find. And also track the product details of customer mostly like, number of users view the product or purchase the product. A reputablebusiness should also have good customer feedback - friends, family orother customers rate them highly.

Supplier

Supplier supplies the product items to multiple stores in a city. And also collects the data details from merchants which product is moving fast and users like mostly. Easily can trackand maintainsupply the demand product to the market by using advance methods like Weighted Frequent Itemset Mining.

Market Research – FIM PERFORMANCE ANALYSIS

WD-FIM (Weighted Downward – Frequent Itemset Mining) algorithmwill be analyzed. All weighted frequent itemsetscan be discovered by WD-FIM algorithm.Each weighted frequent*k*-itemset has at least one weighted frequent (*k -* 1)-itemset.Additionally, each weighted frequent (*k -* 1)-itemset is connected with *CWFIS*1 to get *CWFISk* in WD-FIM algorithm,Consequently, as long asweighted frequent (*k -* 1)-itemsets are complete, weightedfrequent *k*-itemsets are also complete in WD-FIM algorithm.The dataset is scanned for the firsttime to get weighted frequent 1-itemsets.

Algorithm

Apriori algorithm

Apriori algorithm employs an iterative level-wise search for generating frequent item sets. The most significant characteristic of Apriori approach is that it constitutes from the previous frequent item sets rather than all the data items accessed in the transaction when selecting candidate item sets. The frequent item sets refer to the item sets whose supports are greater than or equal to the user's specified minsup. Here, Ck is the candidate item sets, where k is the number of item in the item set. Likewise, Lk represent a k-frequent item set. The Apriori algorithm executes as follows:

1) Ck is generated.

2) Lk is generated from Ck by pruning the item sets.

3) Ck+1 is generated by joining Lk with itself.

WD-FIM (weight judgment downward – Frequent Itemset Mining)

The weight judgment downward closure property and the existence property of weighted frequent subsetsare first proposed and proved. Moreover, the WD-FIM algorithm is proposed. The weighted frequent itemsets are discovered by repeated iteration like U-Apriori algorithm. Obviously, there are significant differences between WD-FIM algorithm and U-Apriori algorithm. First,WD-FIM algorithm is proposed for mining weighted frequentitemsets in uncertain datasets. The basis of the proposed WD-FIM algorithm is theaforementioned weight judgment downward closure propertyand existence property of weighted frequent subsets, butthe downward closure property is used directly to narrowthe searching space of frequent itemsets in U-Apriori algorithm. Based on the aforementioned definitions and theorems,the pseudo code of proposed WD-FIM algorithm is given.The proposed WD-FIM algorithm takes the input as: an uncertain transactional dataset, *DS*; a weight table, *wtable*; a user-specified minimum expected weighted support threshold, ". First of all, the variables like *WFIS* and *WFISk* are initialized. Then it scans the dataset to get the weighted frequent 1-itemsets. Finally, on the basis of the proposed weight judgment downward closure property and the existence property of weighted frequent subsets, the weighted frequent *k*-itemsets will be discovered and all weighted frequent itemsets will be returned

**System Requirements**

# H/W System Configuration

# Processor - Pentium –III

Speed - 1.1 GHz

RAM - 256 MB(min)

Hard Disk - 20 GB

Key Board - Standard Windows Keyboard

Mouse - Two or Three Button Mouse

Monitor - SVGA

# S/W System Configuration

* Operating System :Windows95/98/2000/XP /7
* Application Server : Tomcat5.0/6.X /8.X
* Front End : HTML, Java, Jsp
* Scripts : JavaScript, jquery, ajax
* Server side Script : Java Server Pages.
* Database Connectivity : Mysql.