Journal of
Computer Science & Systems Biology

Vaidya and Deshmukh, J Comput Sci Syst Biol 2015, 8:4
DOI: 10.4172/jcsb.1000194

Research Article Open Access

Predicting Rare Disease of Patient by Using Infrequent Weighted Itemset

Seema Vaidya* and Deshmukh PK

Savitribai Phule Pune University, RSCOE, Pune, India

Abstract

Mining association rule is a key issue in information mining. Nevertheless, the customary models overlook the differences among the trades, and the weighted association rule mining does not process on databases with simply binary attributes. Paper propose a novel frequent patterns and execute a tree (FP-tree) structure, which is an intensified prefix-tree structure for securing compacted, critical information about patterns, and make a capable FP-tree-based mining framework, FP improved function algorithm is utilized, for mining the complete arrangement of patterns by example frequent development. Here in this paper tackles the purpose of making extraordinary and weighted itemsets, i.e. infrequent weighted itemset mining problem. The two novel brilliance measures are proposed for figuring the infrequent weighted itemset mining issue. Moreover, the algorithm are tackled which perform IWI which is more negligible IWI mining. Additionally we used the infrequent itemset for decision based structure. The general problem of the beginning of dependable definite rules is difficult for the grounds that theoretically no provoking procedure without any other person can guarantee the rightness of affected theories. In this manner this system expects the sickness with the exceptional signs. Implementation study shows that proposed algorithm enhances the framework which is effective and adaptable for mining both long and short diagnostics rules. Framework enhances results of foreseeing rare diseases of patient.

Keywords: Association rule; Data mining; IWI mining; Infrequent itemset; Frequent pattern growth

Introduction

Frequent itemsets mining has found expansive used in distinct information mining applications together with clients market-basket examination [1], determination of patterns from page access logs [2], and computation of iceberg-cube [3]. Far coming to investigation has, in this way, been driven in finding capable calculations for general itemset mining, especially in finding association rules [4]. In any case, generally less consideration has been paid to mining of rare itemsets, in spite of the way that it has got basic use in (i) mining of negative association rules from rare itemsets [5], (ii) risk of statistical disclosure assessment where frequent patterns in unknown registration information could be prompt statistical disclosure [6], (iii) fraud detection where uncommon patterns in identified with assessment or finance information may propose unusual development associated with fraudulent behavior [6], and (iv) bioinformatics where uncommon pattern in microarray information may propose hereditary disorders [6]. The expansive group of progressive itemset mining algorithms can be thoroughly requested into two arrangements: (i) candidate creation and-test perfect model and, (ii) pattern growth perfect model. In former studies, it has been appeared that pattern growth based algorithms are computationally speedier on thick datasets. Tests with artificial as well as real-life information display that these algorithms outshine the known algorithms by aspects which range from three for small issues to more than your order of scale for huge issues. We also display how the best functions of the two suggested algorithm can be mixed into a multiple criteria, known as Apriori Hybrid. Scale-up experiments display that Apriori Hybrid machines linearly with the variety of dealings. Apriori Hybrid also has outstanding scale-up qualities with regard to the transaction dimension and the variety of products in the databases. IFP min algorithm moves ahead by changing minimally rare itemsets by dividing the dataset into two segments, one containing a particular thing moreover the other that does not. In case the support vector is excessively high, then less number of progressive itemsets will be delivered achieving loss of gainful association rules. Then again, when the threshold value of support is excessively low, an extensive number of continuous itemsets likewise thus substantial different association rules are made, in this way making it troublesome for the customer to pick the basic ones. Some parts of the issue lies in the fact that a limit is used for making continuous itemsets paying little mind to the length of the itemset. To facilitate this problem, Multiple Level Minimum Support (MLMS) model was proposed [7], where separate breaking points are assigned to itemsets of differing sizes in order to oblige the amount of continuous itemsets mined. This model finds extensive applications in market basket examination [7] for upgrading the amount of association rules delivered. We extend our IFP min calculation to the MLMS framework too.

In particular, we concentrate our interest on two different IWI-support measures: (i) The IWI-support-min evaluate, which depends on a lowest cost operate, i.e., the incident of an itemset in a given deal is calculated by the weight of its least exciting product, (ii) The IWI-support-max measure, which depends on a maximum cost function, i.e., the incident of an itemset in a given deal is calculated by the weight of the most exciting product. Observe that, when working with optimization issues, lowest and highest possible are the most widely used cost functions. Hence, they are considered appropriate for generating the choice of a beneficial part of infrequent weighted data correlations.

Particularly, the following issues have been addressed:

A. IWI and Minimal IWI mining driven by a maximum IWI-support-min threshold, and

B. IWI and Minimal IWI mining driven by a maximum IWI-support-max threshold.

***Corresponding author:** Seema Vaidya, Savitribai Phule Pune University, RSCOE, Pune, India, Tel: 02025696064; E-mail: brahmadhalpe23@gmail.com

Received June 03, 2015; **Accepted** June 22, 2015; **Published** June 24, 2015

Citation: Vaidya S, Deshmukh PK (2015) Predicting Rare Disease of Patient by Using Infrequent Weighted Itemset. J Comput Sci Syst Biol 8: 233-238. doi:10.4172/jcsb.1000194

Task (A) requires finding IWIs and minimal IWIs (MIWIs) such as the item(s) with the least regional attention within each deal. IWIs mined from implementing a highest possible IWI-support-min threshold similar to 180 and their corresponding IWI-support-min principles.

Task (B) requires finding IWIs and MIWIs such as item(s) having maximum regional attention within each deal by taking advantage of the IWI-support-max measure. IWIs excavated by implementing a highest possible IWI-support-max threshold similar to 390. They may signify places of CPUs which contain only under-utilized/idle CPUs at each tested time immediate. Observe that, in this perspective, finding huge CPU mixtures may be considered particularly useful by domain professionals, because they signify huge source places which could be reallocated. To achieve task (A) and (B), we existing two novel methods, namely Irregular Calculated Itemset Miner (IWI Miner) and Minimal Irregular Calculated Itemset Miner (MIWI Miner), which execute IWI and MIWI exploration motivated by IWI-support threshold. IWI Miner and MIWI Miner are FP-Growth-like mining algorithms, whose primary functions may be described as follows: (i) Beginning FP-tree node trimming motivated by the highest possible IWI-support constraint, i.e., early removing of aspect of the look for area thanks to a novel product trimming technique, and (ii) cost function-independence, i.e., they perform in the same way regardless of which constraint (either IWI-support-min or IWI-support-max) is used, (iii) early avoiding of the recursive FP-tree search for in MIWI Miner to prevent getting non-minimal IWIs. The property (ii) creates tasks (A) and (B) comparative, from an algorithmic perspective, provided that an initial information modification phase, which adjusts information loads according to the selected aggregation function, is used before achieving the exploration process.

Starting late, the thought of the research group has moreover been centered on the infrequent itemset mining problem, i.e., discovering itemsets whose reappearance of event in the analyzed information is short of what or proportional to a generally compelling limit. The unverifiable case is that the specific algorithms don't demonstrate comparable patterns to the deterministic case. Case in point, in the deterministic case, the FP-growth algorithm is extraordinary to be an extremely proficient methodology. The basics of a weighted methodology of transactions, i.e., a set of weighted things, are normally evaluated in regards to the relating thing weights. In addition, the guideline itemset quality measures have similarly been customer made to weighted information and used for driving the constant weighted itemset mining system. All past studies on association acknowledge an information illustrate under which transaction get in actuality truths about the things that are contained in every exchange. In various applications, on the other hand, the presence of a thing in an exchange is best gotten by a likelihood measure or probability. As a case, a restorative dataset may contain a table of patient records (tuples), each of which contains an arrangement of symptoms and/or infections that a patient perseveres. Applying affiliation mining on such a dataset licenses us to discover any potential connections among the appearances and infections. By and huge, evidences, being subjective recognitions, would best be addressed by probabilities that show their region in the patients tuples.

In several applications, however, items is not existing or missing in a deal, but rather the prospect of it being in the deal is given. This is the case for information gathered from trial dimensions vulnerable to noise. For example, in satellite picture data the existence of an objector function can be indicated more consistently by a probability score when it is acquired by very subjective human interpretation or an image segmentation device. In frequent itemset mining, the regarded deal dataset is generally showed as a binary matrix M where each line symbolizes a deal and every line matches to products. Mining frequent patterns from this type of datasets is more difficult than exploration from conventional deal datasets. After all, processing the assistance of an itemset now has to take the lifestyle likelihood of the products into consideration. To provide information about the regularity of an itemset, two techniques are available.

Existing System

The problem of mining frequent itemsets was at first represent by Grahne [8], who proposed the Apriori algorithm. Apriori is a methodology of bottom up, BFS (Breadth First Search) algorithm that attempts the decreasing conclusion property "all subsets of frequent itemsets are successive". Candidate frequent itemsets just whose subsets are all frequent are made in every database filter. Apriori needs l database checks on the off chance that the extent of the greatest regular itemset is l.

Efficient algorithms for mining regular itemsets are crucial for mining association rules as well as for many other information mining tasks. Means of exploration regular itemsets have been applied using a prefix-tree framework, known as an FP-tree, for saving compacted information about regular itemsets. Numerous trial outcomes have confirmed that these techniques perform extremely well. In paper [8], we current a novel FP-tree strategy that greatly decreases the need to navigate FP-trees, thus obtaining significantly improved performance for FP-tree-based algorithms. Our technique works especially well for rare information set. Furthermore, we current new algorithms for exploration all, maximum, and shut regular itemsets. Our algorithms use the FP-tree information framework in combination with the FP-array strategy efficiently and incorporate various optimization techniques. We also carried out current trial outcomes evaluating our techniques with current algorithms. The outcomes show that our techniques are the quickest for many cases. Even though the algorithms eat much storage when the information places are rare, they are still the quickest ones when the minimum support is low. Moreover, they are always among the quickest algorithms and eat less storage than other algorithms when the information set are heavy.

Han et al. presented new algorithm known as the FP-growth methodology for mining frequent itemsets. The FP-growth methodology is a DFS algorithm [9]. A data structure called the FP-tree is used for removing the repeat information of itemsets in the first exchange database in a compacted structure. Only two database checks are needed for the algorithm, and no candidate generation is required. This makes the FP-growth technique much faster than Apriori. A lot of new issues may happen when we at the same time research positive and negative association rules (PNARs). These issues consist of how to find irregular itemsets, how to produce PNARs properly, how to fix the problem due to a single lowest assistance and so on. Infrequent itemsets become very important because there are many respected adverse negative association rules (NARs) in them. In our past work, a MLMS design was suggested to find at the same time both regular and irregular itemsets by using several level lowest facilitates (MLMS) design. In this document, a new evaluate VARCC which brings together connection coefficient and lowest assurance is suggested and a corresponding criteria PNAR_MLMS is also suggested to produce PNARs properly from the regular and irregular itemsets found by the MLMS design. The trial results show that the evaluate and the algorithm are effective.

Grahne et al. introduced a novel FP-array technique that phenomenally decreases the need to cross the FP-trees [6]. This paper explains each of these stages in details. Given its application potential, Web utilization exploration has seen a fast improve in attention, from both the analysis and practice communities. This document provides a specific taxonomy of the perform in this place, such as analysis efforts as well as professional offerings. An up-to-date study of the current perform is also offered. Lastly, a brief summary of the Web SIFT program as an example of a prototypical Web utilization exploration program is given.

Chun Kit Chui, the issue of mining frequent itemsets from uncertain data under a probabilistic structure [3]. We consider transactions whose products are associated with existential possibilities and provide a formal definition of regular styles under such an unclear information design. We display that traditional algorithms for mining regular itemsets are either inapplicable or computationally ineffective under such a design. An information cutting structure is suggested to enhance exploration performance. Through comprehensive tests, we display that the data trimming technique can accomplish important benefits in both CPU price and I/O price.

There is a significant collection of research on Frequent Itemset Mining (FIM) however no road numbers FIM in unverifiable databases [5,9,10]. The system proposed by Chui et al. forms the normal support of itemsets by total summing all itemset probabilities in their U-Apriori algorithm [5]. Later, they besides proposed a probabilistic later with a particular end objective to prune candidates early [5]. The UF-growth algorithm is proposed [10]. Like U-Apriori, UF-growth figures frequent itemsets by system for the normal support; nonetheless it uses the FP-tree approach to keep up a key separation generation of candidates [8]. Instead of our probabilistic strategy, itemsets are viewed incessant if the expected support surpasses minsup. The major drawback of this estimator is that information about the vulnerability of the expected support is lost; [2,10,11] neglect the amount of possible things in which an itemsets is frequent; proposes right and sampling algorithms to discover likely frequent things in streaming probabilistic data [12]. In any case, they don't consider itemsets with more than one thing. Finally, except for [13], existing FIM algorithm expect parallel esteemed things which blocks fundamental adjustment to unverifiable databases. Existing systems in the field of uncertain data administration and mining can be arranged into different exploration headings. Most related to our work are the two arrangements probabilistic databases" [4,11,14,15] and probabilistic query transformation" [6,10,16,17]. The vulnerability model used as a part of our system is near to the model used for probabilistic databases.

The Iceberg-Cube problem reduces the calculations of the data cube to only those group-by categories fulfilling a lowest threshold condition described on a specified evaluate. In paper [3], we apply the Bottom-Up Computation (BUC) algorithm for processing Iceberg cubes and perform a understanding research of BUC based on the possibility solidity operate of the information. The withdrawals under consideration are the Gaussian, Geometric, and Poisson distributions. The Consistent submission is used as a foundation for evaluation. Results show that when the cube is sparse there is a connection between the information submission and the running time of the algorithm. In particular, BUC works better on Uniform followed by Poisson, Gaussian and Geometrical information.

A probabilistic database demonstrates a database made out of relations with questionable tuples [6], where each tuple is associated with likelihood significance the probability that it exists in the relations. This model, called tuple vulnerability", gets the conceivable planets semantics [4]. A probabilistic database represents an arrangement of possible certain" database amples (planets), where a database example identifies with a subset of uncertain tuples. Every one example (world) is joined with the probability that the world is valid". The probabilities mirror the likelihood dissemination of all possible database occurrences. In the general model depiction [15], the possible things are obliged by concludes that are on the tuples remembering the finished objective to unite object (tuple) connections. The ULDB model proposed in [14], which is used as a part of Trio [1], supports unverifiable tuples with alternative occurrences which are called x-relations. Relations in ULDB are called x-relations containing an arrangement of x-tuples. Each x-tuple contrasts with an arrangement of tuple events which are thought to be regularly world class, i.e. near to one event of an x-tuple can appear in a possible world event in the meantime. Probabilistic top-k query approaches [11,17] are normally connected with databases which are uncertain using the tuple weakness model. The technique proposed in [17] was the first approach prepared to enlighten probabilistic questions effectively under tuple in dependence by strategy for programming methodologies which is dynamic.

Frequent itemset mining is a basically used data mining framework that has been introduced in [1]. In the traditional method of itemset mining issue things having a place with value based data are managed also. To allow differentiating things considering their point of interest or power inside every exchange, in [14] the analysts focus on discovering more valuable association rules, i.e., the weighted association guidelines (WAR), which incorporate weights importance thing criticalness. The problem of mining association rules over container information was presented in [4]. An example of such a concept might be that 98% of clients that purchase wheels and automatic components also get automobile services done. Discovering all such guidelines is useful for combination marketing and connected emailing programs. Other programs consist of catalog design, add-on sales, store layout, and client segmentation based on purchasing patterns. The information sources engaged in these programs are very large. It is crucial, therefore, to have fast algorithms for this process. Regardless, weights are displayed just in the midst of the generation of rule performing the routine frequent itemset mining procedure. The principle attempt to pushing thing weights into the itemset mining procedure has been done in [7]. We also introduce the novel concept of residual trees. We further utilize the residual trees to mine multiple level minimum support itemsets where different thresholds are used for finding frequent itemsets for different lengths of the itemset. Finally, we analyze the behavior of our algorithm with respect to different parameters and show through experiments that it outperforms the competing ones. It proposes to try the resistance to monotonicity of the proposed weighted support basic to drive the Apriori-based itemset mining stage. Regardless, in [7,14] weights need to be Pre-assigned, while, in various honest to goodness cases, this may not be the circumstance. To address this issue, in [15] the analyzed quality based data set is talked to as a bipartite center power diagram and surveyed by technique for a remarkable indexing system, i.e., HITS [3], with a particular deciding objective to robotize thing weight undertaking. Weighted thing backing and sureness quality records are described as requirements be also, used for driving the itemset and guideline mining stages. This paper contrasts from the strategies on the grounds that it focuses on mining infrequent itemsets from weighted data of frequent ones. Subsequently, different pruning procedures are abused.

In this research, we recommend a novel frequent-pattern shrub (FP-tree) framework, which is a prolonged prefix-tree framework for saving compacted, crucial information about regular styles, and develop an effective FP-tree, based mining technique, FP-growth, for exploration the complete set of regular styles by design fragment development. Performance of exploration is obtained with three techniques: (1) A huge database is compacted into a compacted, more compact data framework, FP-tree which prevents expensive, recurring database tests, (2) our FP-tree-based exploration assumes a pattern-fragment growth technique to avoid the expensive generation of a huge number of candidate sets, and (3) a partitioning-based, divide-and-conquer technique is used to break down the exploration task into a set of more palatable pieces for exploration limited styles in depending data source, which considerably decreases the search space. Our performance research shows that the FP-growth technique is effective and scalable for exploration both long and short regular styles, and is about an order of scale quicker than the Apriori algorithm and also quicker than some recently revealed new frequent-pattern exploration methods.

A related exploration issue is probabilistic regular itemset mining [5,15]. It includes mining frequent itemsets from unverifiable data, in which thing occasions in every trade are unverifiable. To address this issue, probabilistic models have been created and facilitated in Apriori-based [5] or projection-based [11] calculations. In spite of the way that probabilities of thing occasion may be remapped to weights, the semantics behind probabilistic and weighted itemset mining is generally assorted. Without a doubt, the probability of occasion of a thing inside a trade may be totally uncorrelated with its relative hugeness. For instance, a thing that is at risk to happen in a given trade may be regarded the least essential one by a master within domain.

Carson Kai-Sang Leung, mining of frequent patterns is one of the popular knowledge finding and data mining (KDD) tasks [5]. It also performs an essential part in the mining of many other styles such as correlation, sequences, and association rules. Hence, it has been the topic of numerous researches since its release. Most of these researches discover all the frequent patterns from selection of accurate information, in which the items within each datum or deal are definitely known and accurate. However, there are many real-life circumstances in which the customer is interested in only some small areas these regular styles. Finding all frequent patterns would then be repetitive and waste lots of calculations. This demands restricted mining, which is designed to discover only those regular styles that are exciting to the customer. Moreover, there are also many real-life circumstances in which the information are unclear. This demands unclear information mining. In this article, we recommend methods to effectively discover restricted regular styles from selections of unclear information.

In paper, presents U-relations, a helpful and simply relational reflection system for unclear databases. U-relations support attribute-level doubt using straight dividing [16]. If we consider positive relational algebra prolonged by a function for handling possible solutions, a query on the sensible stage can be converted into, and analyzed as, a single relational geometry question on the U-relation reflection. The interpretation plan basically maintains the size of the question with regards to variety of functions and, in particular, variety of connects. Standard techniques employed in off-the-shelf relational data source management systems are effective for improving and handling concerns on U-relations. In our tests we show that question assessment on U-relations machines to considerable amounts of data with high levels of uncertainty.

Probabilistic regular itemset mining in unclear deal data source semantically and computationally varies from conventional methods used to conventional "certain" deal databases. The concern of existential doubt of item (sets), showing the possibility that an item (set) happens in a deal, creates conventional methods inapplicable. In [17], we existing new probabilistic remedies of regular itemsets depending on possible world semantics. In this probabilistic perspective, an itemset X is known as frequent if the possibility that X happens in at least minSup dealings is above a given threshold τ. To the best of our information, this is the first strategy dealing with this issue under possible world's semantics. In concern of the probabilistic remedies, we existing a framework which is able to fix the Probabilistic Frequent Itemset Mining (PFIM) issue effectively. A comprehensive trial assessment looks into the effect of our suggested methods and reveals that our strategy is purchases of scale quicker than straight-forward techniques.

Proposed System

Proposed system

Following Figure 1 shows the proposed system architecture. First it gives dataset (patient's symptoms) for mining preprocess. Algorithm retrieves infrequent itemset through dataset. After taking infrequent itemset, algorithm sorts out minimal infrequent itemset using threshold. So, we detect any rare disease from the minimal infrequent itemset.

Function enhanced FP algorithm: The advantage of function enhanced FP algorithm is that it works without any tree or any other complex data structure. The steps of the algorithm are as follows:

Step 1: Initially evaluate the support of the items.

Step 2: The initial transaction database is converted in to a set of transaction list, with one list for each item.

Step3: The first list corresponding to the item e contains the second, seventh and eight transactions, with the item e, removed.

Mining process: After implementing the function enhanced FP algorithm. The rules are generating, by using the enhanced FP algorithm. Process of mining is done, now we implement IWI and MIWI mining process.

IWI mining

IWI mining is similar to FP-growth mining algorithm. It performs the step similar to FP-growth mining algorithm. The steps are as follows:

Step 1: Initially the FP tree is generated.

Step 2: recursive itemset mining from the FPtree index.

The main difference is that this algorithm generated rare weighted

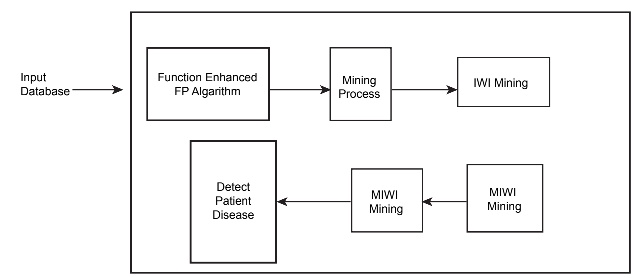

Figure 1: Framework of the proposed system.

rules instead of weighted rules. For generating the rare rules following steps are as follows:

Step 1: A new pruning method for pruning the rules is proposed.

Step 2: A slightly modified FP tree structure is used for storing the IWI support value associated with each node.

MIWI mining process: Given a weighted transactional data set and a maximum IWI-support threshold, the Minimal Infrequent Weighted Itemset Miner algorithm extracts all the MIWIs that satisfy the minimal threshold.

Algorithms

```
Function Enhanced-FP (a: array of transaction lists,
p: set of items,
s min: int): int
Var i; k: item;
s: int;
n: int;
b: array of transaction lists;
t; u: transaction list element;
begin
n: = 0;
While a is not empty do
i: = last item of a; s := a[i].wgt;
If s >= smin then
p: = p È{i};
Report p with support s;
p: = p – {i};
end;
If s >= smin then
b: = array of transaction lists;
t: = a[i].head;
While t = nil do
u: = copy of t; t: = t.succ;
k: = u.items [0];
remove k from u.items;
If u.items is not empty
then u.succ = b[k].head; b[k].head = u; end;
b [k].wgt:= b[k].wgt u.wgt;
end;
n: = n + 1 + Enhanced-FP (b; p; smin);
end;
t: = a[i].head;
while t=nil do
u: = t; t: = t.succ;
k: = u.items [0]
Remove k from u.items;
if u.items is not empty
then u.succ = a[k].head;
a[k].head = u; end;a[k].wgt:= a[k].wgt + u.wgt;
end; remove a[i] from a;
end;
return n;
end;
```

Experimantal Expected Result

Dataset

In this work we have used patient dataset for predicting diseases. For this we have specially used diabetes dataset. Firstly we have to process dataset, to convert it into required format. This dataset is taken from http://repository.seasr.org/Datasets/UCI/arff/diabetes.arff

Results

In the above Table 1 it shows comparative analysis of existing system and proposed system. For every different transaction the time required in proposed system is always less than existing system. So, our experimental evaluation results shows that proposed system improves accuracy in terms of time requirement. Following Figure 2 shows graph for average response time according to Table 1.

Sl. No.	Average response time in seconds		
	Number of transaction	Existing system	Proposed system
1	2000	2.6	2.4
2	4000	3	2.6
3	6000	3.2	2.8
4	8000	4	3.3
5	10000	4.2	3.8

Table 1: Comparative analysis.

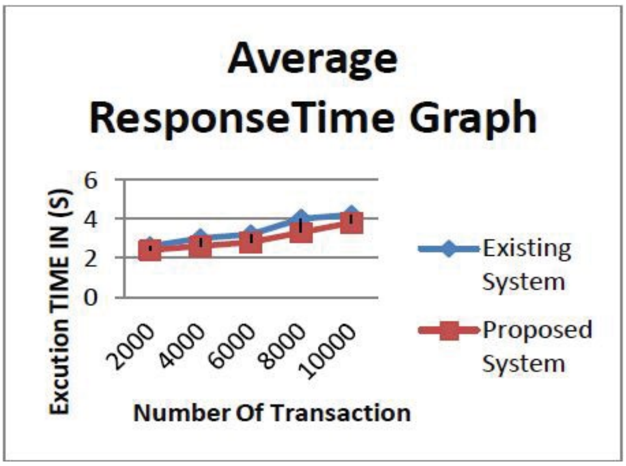

Figure 2: Average response time graph.

Conclusion

In this paper we introduce the methodology of infrequent itemset mining from frequent patterns. In our methodology we are anticipating diseases of patients by utilizing rare weighted itemsets. Experimental assessment results demonstrate that our proposed framework enhances result than existing framework. In this framework, function enhanced FP algorithm is utilized which is readily helps to enhance results. This algorithm helps to a more noteworthy degree. In future, we can upgrade our framework with cutting edge choice making framework. In this way, it can support to the target activity depend on features of found IWIs.

References

1. Agrawal R, Srikant R (1994) Fast algorithms for mining association rules in large databases. Proceedings of the 20th VLDB Conference, pp. 487-499.
2. Chui CK, Ben Kao, Hung E (2007) Mining frequent itemsets from uncertain data. In 11th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining 4426: 47-58.
3. Beyer K, Ramakrishnan R (1999) Bottom-up computation of sparse and iceberg cube. SIGMOD Rec USA 28: 359-370.
4. Dong X, Niu Z, Shi X, Zhang X, Zhu D (2007) Mining both positive and negative association rules from frequent and infrequent itemsets. ADMA 4632: 122-133.
5. Kai-Sang Leung C, Christopher L Carmichael, Boyu Hao (2007) Efficient mining of frequent patterns from uncertain data. In ICDMW '07: Proceedings of the Seventh IEEE International Conference on Data Mining Workshops, USA. Pp-489-494.
6. Agrawal R, Srikant R (1994) Fast algorithms for mining association rules in large databases. VLDB, pp. 487-499.
7. Haglin DJ, Manning AM (2007) On minimal infrequent itemset mining. DMIN. Pp. 141-147.
8. Grahne G, Zhu J (2005) Fast algorithms for frequent itemset mining using fp-trees. Trans Know Data Engg 17: 1347-1362.
9. Han J, Pei J, Yin Y (2000) Mining frequent patterns without candidate generation. SIGMOD '00 Proceedings of the 2000 ACM SIGMOD international conference on Management of data. Pp.1-12.
10. Chun Kit Chui and Ben Kao (2008) A decremental approach for mining frequent itemsets from uncertain data. In The 12th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD) 5012: 64-75.
11. Qin Zhang, Feifei Li, Ke Yi (2008) Finding frequent items in probabilistic data. In Jason Tsong-Li Wang, editor, SIGMOD Conference. Pp. 819-832.
12. Srivastava J, Cooley R (2000) Web usage mining: Discovery and applications of usage patterns from web data. SIGKDD Exploration 1: 12-23.
13. Florian Verhein, Sanjay Chawla (2006) Geometrically inspired itemset mining. In IEEE International Conference on Data Mining (ICDM 2006), IEEE Computer Society, pp. 655-666.
14. Dong X, Zheng Z, Niu Z (2007) Mining infrequent itemsets based on multiple level minimum supports. ICICIC, pp. 528.
15. Jiawei Han, Jian Pei, Yiwen Yin (2000) Mining frequent patterns without candidate generation. SIGMOD Rec 29: 1-12.
16. Antova L, Jansen T, Koch C, Olteanu D (2008) Fast and Simple Relational Processing of Uncertain Data. In Proc. 24th Int. Conf. on Data Engineering (ICDE'08), Cancun, Mexico.
17. Bernecker T, Kriegel HP, Renz M, Verhein F, Zuefle A (2009) Probabilistic Frequent Itemset Mining in Uncertain Databases, Proc. 15th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD '09). Pp. 119-128.