Survey on techniques of High Utility Mining

Ms. Archana Kisan Dere [1], Prof. S. A. Kahate[2]

[1] *Student of M.E. Second Year, Computer Engineering*, SPCOE , Dumberwadi ,Otur , *Maharashtra, India*

[2] *Assistant Professor, Computer Engineering*, SPCOE , Dumberwadi , Otur , *Maharashtra, India*

ABSTRACT

*Frequent Itemset mining is essential in data mining and its popular application is market basket analysis. Traditional FIM model produce large number of frequent but low revenue itemset and lose information on valuable itemset with low selling frequencies. Utility mining is essential to avoid this. Discovery of high utility itemset (e.g. high profit,weight) from database is an important task. However, discovered HUIs may not be compact in size which degrades efficiency of mining process. To achieve efficiency, some compact representations available such as free sets, non-derivable sets and closed itemsets. Mining closed high utility itemsets serves as a compact and lossless representation of HUIs. To achieve this, algorithm CHUD (Closed High Utility Itemset Discovery) is proposed in literature. Organizations mine huge data to generate HUI. Organizations conducting business have to suffer the huge data mining computation and storage overhead. To avoid the computation burden of mining enormous data, there is a scope to outsource this task to another party. But outsourced party may not be trustworthy so there's the need of preservation of data privacy from outsourced party. Various privacy preservation techniques are available in literature. And instead of mining whole database again for getting HUIs, by using techniques from [10] we can also recover all HUI's from the concise representation. In this paper we have studied different techniques available for the compact representation of HUIs***.**

Keyword - *Frequent Itemset mining, Closed High Utility Itemset , Utility mining*

I. INTRODUCTION

Data Mining comprises mining huge amount of data using different mining techniques to derive useful and important information i.e. discovery of knowledge. Organizations use such knowledge for their respective purposes. Frequent Itemset Mining (FIM) is famous technique for market basket analysis i.e. discovery of itemsets that customer purchase together frequently. But some problems are present in FIM model that it generates high amount of frequent itemsets and avoids low selling frequencies. Generated itemsets may have low revenues and less frequent itemsets may generate high revenue. FIM treats every item with same importance, profit and weight and it assumes an item can be either present or absent i.e. binary representation of itemsets in transaction. To overcome this problem, utility mining concept is emerged. A utility of an itesmset is measured in cost, weight, quantity, profit etc. If the utility of an itemset is greater than user-specified minimum utility threshold then it is said to be high utility itemset otherwise it is considered as a low utility itemset.

High utility mining has wide application such as website click stream analysis, mobile commerce environment ,biomedical applications etc. HUI may discover large amount of high utility itemsets causes difficulty to user for result analysis. It also requires more memory and processing time tends to be less efficient. To reduce cost overhead and mining task and to provide concise representation, different approaches like Freeset[6], Non derivable sets[1], Closed Itemset[5] are introduced in FIM. But applying these techniques to HUI produced several challenges:

1.Lossy representation of all HUI which is not meaningful to user;
2. Significant reduction in the extracted patterns may not be achieved;
3. Algorithms may not be efficient;
4. How to recover all HUI's from the concise representation.

Algorithms proposed in [20] integrate the concept of closed itemset into high utility itemsets mining to address mentioned challenges and named iii) Closed+ High Utility itemset Discovery (CHUD). Also few strategies were incorporated for performance enhancement. Storage of all itemsets and their utility information requires excessive memory and processing them to get set of CHUIs is of highly computational task. Organizations conducting business have to suffer the huge data mining computation and storage overhead. Organizations have huge data and tend to outsource the task of HUI mining to another party. In existing systems, the outsourced party will have the information about High Utility Items and their utility (e.g. Profit) so does not preserve the data privacy. Our proposed systemis to alleviate the load of computation, storage and processing to another property with preservation of privacy of outsourced high utility mining and also to produce the concise representation of HUIs using existing work.

Organizations need to mine huge amount of data to generate HUI. Organizations conducting business have to suffer the huge data mining computation and storage overhead. To avoid the computation burden of mining enormous data, we introduced to outsource this task to another party. But there is cas e when outsourced party may not be fully trusted. So there is a need of preservation of data privacy from outsourced party. We used technique to preserve the data privacy from outsourced party. To achieve the privacy in such scenario, we used one to one su bstitution technique which will converts the items utility into some random numbers and mapping of these associations is stored locally. To perform operations after this substitution we employed homomorphic encryption technique based on pallier cryptosystem. And instead of mining whole database again for getting HUIs, by using techniques from [20] we can also reconstruct all HUI's from the concise representation.

2. LITERATURE REVIEW

R. Agrawal et.al[1]considered the problem of discovering association rules between items in a large database of sales transactions and presented two new algorithms for solving this problem that are fundamentally different from the known algorithms and also showed how the best features of the two proposed algorithms can be combined into a hybrid algorithm called Apriori Hybrid Scale up. Experiments showed that Apriori Hybrid scales linearly with the number of transactions and also has excellent scale up properties with respect to the transaction size and the number of items in the database.

R.Agarwal et.al [2] In database mining, association rules discovery is an important problem. R.Agarwal et.al. , introduced new algorithms for fast association mining, which scan the database only once, eliminating the need of multiple database scans. The algorithms used novel itemset clustering techniques to approximate the set of potentially maximal frequent itemsets. The algorithms then make use of efficient lattice traversal techniques to generate the frequent itemsets contained in each cluster. Paper proposed two clustering schemes based on equivalence classes and maximal hyper graph cliques, and study two traversal techniques based on bottom-up and hybrid search .Experiments showed improvements in association rule discovery process.

N. Pasquier et.al . [3] Proposed closed algorithm because in data mining, discovery of association rules is a most important task. Identifying the relationships between the items in larger databases is the main objective of discovering the association rules. For this many efficient algorithms are evolved like Mannila's algorithm, Partition, sampling all are based on Apriori mining method i.e. subset lattice pruning. Based on a new mining method i.e. pruning the closed set lattice, paper proposed a Close algorithm. The algorithm is efficient and optimized version of Apriori. Close algorithm proves to be efficient for mining dense and correlated data.

Motwani, R et.al.[4] Mining association rules is the most interesting issues in the data mining area. Dynamic Miss-Counting algorithms find association rules with confidence pruning but without support pruning. To handle data sets with a large number of columns, this paper proposed dynamic pruning techniques that can be applied during data scanning. DMC counts the numbers of rows in which each pair of columns disagree instead of counting the number of hits. DMC deletes a candidate as soon as the number of misses exceeds the maximum number of misses allowed for that pair. Authors also introduced several optimization techniques that reduce the required memory size significantly, thus improving overall performance.

Jian Pei et.al[5]In frequent pattern mining, due to numerous problems like, huge set of candidate sequence generation, number of repeated scans of the database and problem of Apriori while mining long sequential patterns, this paper proposed an efficient sequential pattern mining method named as Frequent pattern projected Sequential pattern mining i.e. FreeSpan. It integrates the mining of frequent sequences with frequent mining and in order to confine the search space used projected sequence database and it reduces the efforts of candidate subsequent generation.

T. Calders et.al [6] investigated Frequent Itemset mining was popular application to generate frequently purchased itemsets in market basket analysis. But it can generate high amount of frequent itemsets if the data is highly correlated and set minimum support threshold is very low. Instead of mining all frequent items the solution was to construct concise representation of frequent itemsets. In this Paper their aim was to identify the redundancy of frequent itemsets to reduce the result of mining operation. Paper presents the deduction rules allows the minimal representation of all frequent itemsets. Non Derivable Itemsets considered for concise representation. Experiments show that mining concise representation first and then from this creating frequent itemsets give better results than existing algorithms.

J.F. Boulicaut et.al [7] investigated extracting the frequent itemset efficiently in a large collection of transactions containing items is a common data mining problem. Frequent Itemset Mining produces high amount of data. Paper proposed a structure called free-sets, which provides the base for approximation of any itemset support i.e. the number of transaction containing the itemset. A new ε-adequate representation for the frequency queries is introduced. This representation, called free-sets, is more concise than the ε-adequate representation based on itemsets by Mannila and Toivonen. Using pruning strategies developed for frequent itmeset discovery, frequent free-sets can be easily extracted which can be further used to calculate the support of any frequent itemsets and this support is used to approximate very closely the support of frequent itemsets. Then, paper considers this approximation on association rules i.e. patterns derived from frequent itemsets.

W. Cheung et.al [8] introduces Compressed and Arranged Transaction Sequences Tree or CATS Tree and CATS Tree algorithms. Once CATS Tree is built, it can be used for multiple frequent pattern mining with different supports. Furthermore, CATS Tree and CATS Tree algorithms allow single pass frequent pattern mining and transaction stream mining. In addition, transactions can be added to or removed from the tree at any time. CATS Tree extends the idea of FP-Tree to improve storage compression and allow frequent pattern mining without generation of candidate itemsets.

S.F. Shie et.al. [9] Some algorithms, for e.g. Apriori, suffer from the costs to handle a huge number of candidate sets and scan the database repeatedly. An algorithm, named FP-growth, for mining the complete set of frequent itemsets from FP-tree is developed. This approach avoids the costly generation of a large number of candidate sets and repeated database scans, which was regarded as the most efficient strategy for mining frequent itemsets. Updates to the transaction database could invalidate existing rules or introduce new rules. The problem of updating the association rules can be reduced to finding the new set of frequent itemsets in the updated database. A simple solution to the update problem was to re-mine the frequent itemsets of the whole updated database. However, it was clearly inefficient because all the computations done in the previous mining are wasted. This problem is avoided in proposed algorithm, called AFPIM(Adjusting FP-tree for Incremental Mining), to efficiently find new frequent itemsets with minimum re-computation.

R. C.W. Wong et.al.[10] considers the scenario of data stream. Data is continuously arriving through the data stream in terms of sliding window. To mine the K most interesting itemsets of varied sizes in a data stream the authors adopted new model using the sliding window concept. Sliding window is partitioned into buckets. For the transactions in each bucket, statistics of the frequency counts of the itemsets is maintained. Experimental studies showed that algorithm guarantees no false negatives for any data distributions and also show that the number of false positives returned is typically small according to Zipfian Distribution. Experiments on synthetic data showed that the memory used is tens of times smaller than that of a naive approach, and the false positives are negligible.

C. Lucchese et.al .[11] Proposed High Utility Pattern mining had been several applications in broader aspect. Efficient mining of high utility itemsets was very important in data mining . C. Lucchese et.al. proposed two one-pass algorithms named MHUI-BIT and MHUI-TID for mining high utility itemsets from data streams within transaction-sensitive sliding window. To improve efficiency of HUI, Two effective representations of item information and an extended lexicographical tree-based summary data structure are developed. Proposed algorithms provide better results than existing algorithms used for HUI mining from data streams.

U. Yun et.al[12]To prune the search spaces i.e. removing the itemsets which are not satisfying certain condition many algorithms used a support constraint. This scheme allows for basic pruning but the resulting patterns have weak affinity after mining datasets for obtaining frequent patterns. U. Yun et.al , proposed an efficient mining algorithm called WIP (Weighted Interesting Pattern mining). The algorithm founded on mining weighted frequent patterns. This paper determine the concept of a weighted hyperclique pattern that uses a new measure, called weight-confidence, to consider weight affinity and prevent the generation of patterns with substantially different weight levels. Experimental study showed that WIP is efficient in weighted frequent pattern mining and it generates less but more worthful patterns for users.

Erwin et.al[13] used pattern growth approach (Han, Wang, & Yin, 2000) and proposed an efficient algorithm for utility mining. Revent research by authors (Erwin, Gopalan, & Achuthan, 2007) on utility mining had produced an algorithm that ran efficiently on dense data, but performed unsatisfactorily on sparse data. They felt the need to improve the performance on sparse data, motivated by this they have formulated a new compact data representation named Compressed Utility Pattern tree (CUP-tree) which extends the CFP-tree of (Y.G Sucahyo & R.P Gopalan, 2004) for utility mining, and a new algorithm named CTU-PRO for mining the complete set of high utility itemsets.

A. Erwin et.al [14]proposed an algorithm that uses Transaction Weighted Utility with pattern growth based on a compact utility pattern tree data structure and algorithm implements a parallel projection scheme to use disk storage when the main memory is inadequate for dealing with large datasets. Experimental evaluation showed that the algorithm is more efficient compared to previous algorithms and can mine larger datasets of both dense and sparse data containing long patterns.

K. Chuang et.al[15]explores a practicably interesting mining task to retrieve top-k (closed) itemsets in the presence of the memory constraint and attempts to specify the available upper memory size that can be utilized by mining frequent itemsets. To comply with the upper bound of the memory consumption, two efficient algorithms, called MTK and MTK Close, are devised for mining frequent itemsets and closed itemsets, respectively, without specifying the subtle minimum support, both achieved high efficiency and have a constrained memory bound, showing the prominent advantage to be practicable algorithms of mining frequent patterns.

Y.C. Li et.al.[16]proposed the Isolated Items Discarding Strategy (IIDS), which can be applied to any existing level-wise utility mining method to reduce candidates and to improve performance. The most efficient known models for share mining are ShFSM and DCG, which also work adequately for utility mining as well. By applying IIDS to ShFSM and DCG, the two methods FUM and DCG+ were implemented, respectively.

B. Le.et.al.[17] explored to apply fuzzy sets theory to the utility mining problem and propose a novel method, namely *FHUI (Fuzzy High Utility Itemsets)-Mine*, for mining fuzzy high utility itemsets. In addition to reflect the fuzzy degree for quantity and profit regions of high utility itemsets, *FHUI-Mine* also provides a fuzzy threshold range that may include itemsets with profits slightly less than the designated threshold value.

C. F. Ahmed et.al. [18] Developed the novel tree structure because High Utility Pattern extraction has high value as it considers different profit values for every item and non-binary frequency values of items in transaction. Cost overhead increased when the database is updated or minimum threshold is changed. In order to reduce these unnecessary calculations, incremental and interactive data mining provides ability to use previous mining results and data structures. This paper integrates the HUP and incremental and interactive mining and proposed three novel tree structures to perform interactive and interactive HUP mining more efficiently. Incremental HUP lexicographic order is a first tree structure arranged as per item's lexicographic order. It has ability to pick up incremental data without need of any restructuring operation. To obtain compact size by arranging items as per their transaction frequency, second tree structure IHUP Transaction Frequency Tree is designed. The third tree structure IHUP Transaction-Weighted Utilization is designed on the basis of TWU value of items to reduce the mining time. The mechanism provides efficiency and scalability in incremental and interactive HUP mining.

Engelbert Mephu Nguifo et.al. [19] was introduced exact concise representation based on discovery of the disjunctive search space. The di1sjunctive itemsets are able to deliver the information about the complementary occurrences of items residing in the dataset. In the disjunctive search space, disjunctive supports are used to identify the respective itemsets. But there is possibility of a redundancy because itemsets may characterize the same data and this is not self-contained process. This can be avoided by the use of closure operator to get the compact and concise representation. This paper focused on new closure operator related to the disjunctive search space, disjunctive and negative supports are considered. Approach aimed at retaining disjunctive closed itemsets and ensuring concise and exact recovery of the frequent itemsets with respect to minimum length description principle. Experiments carried out showed the technique as effective.

B. Vo et.al.[20] , present a parallel method for mining HUIs in vertically partitioned distributed databases. Use WIT-tree structure to store local database on each site for parallel mining HUIs. The item ith in each SlaverSite is only sent to MasterSite if its Transaction-Weighted Utilization (TWU) satisfies minutility (minutil), and MasterSite only mines HUIs which exist at least on 2 sites. Besides, the parallel performance is also interesting because it reduces the waiting time of attended sites. Thus, the mining time is reduced more significant than that in mining from centralized database.

V. S. Tseng et.al.[21] Mining high utility itemsets from large database may generate large number of candidate iemsets. This leads to degradation of mining performance in terms of execution time and space requirement. The performance goes worse when there is huge number of long transactions or long high utility itemsets. V. S. Tseng et.al. , proposed an efficient algorithm, called UP-Growth (Utility Pattern Growth), for mining high utility itemsets with a set of techniques for pruning candidate itemsets. The information of high utility itemsets is maintained in a special data structure named UP-Tree (Utility Pattern Tree) such that the candidate itemsets can be generated efficiently with only two scans of the database. Results showed that the algorithm reduces number of candidates.

B.E. Shie et.al.[22] proposed mobile sequential patterns i.e. moving path sequences with transactions. In this paper they has been combined high utility mining with the mobile data mining and proposed mining high utility mobile sequential patterns approach. To achieve this,they proposed two tree based algorithm using different strategies using breadth first and depth first generation techniques. Experiments showed that algorithm performed better than existing mobile sequential pattern algorithms.

T. Hamrouni et.al. [24] concentrated on exact concise representations of frequent patterns and described their close relation with important concepts like the framework of ε-adequate representation and the minimum description length principle. Based on the mathematical settings of Formal Concept Analysis, it showed the complementarity between minimal generators and closed itemsets. Then, the paper focused on the key role played by these patterns for solving several problems associated to various pattern classes. In this respect, it classified concise representations of frequent itemsets according to their common characteristics.

Mengchi Liu et.al. [25] Proposed the HUI algorithm because process of mining high utility itemset mining gives valuable information to the businesses. For the generation of HUI , most of the algorithms firstly estimate the utilities of the itemsets and generate the candidate itemsets and then by scanning the database compute the exact utilities of the itemsets to generate the high utility itemsets. Most of the existing algorithms tend to generate large number of candidate itemsets and suffers with problems like: Storage of large number of candidate requires large memory, a lot of time consumed for candidate generation and their exact utility computation which results in the performance degradation and lesser efficiency. In this Paper they was proposed an algorithm for mining HUIs without generation of the candidates. Proposed a structure called utility-list which stores the utility information about the itemsets and also stores the heuristics information about the decision of pruning. An algorithm named as, HUI-Miner is able to mine high utility itemsets from the initially constructed utility-list. This approach is compared with various algorithms like IHUPTWU], UPGrowth etc. Results showed that HUI-Miner shows significant improvement in the performance than these algorithms.

K. Subramanian et.al.[26] For mining high utility itemsets organizations focus on centralized data location where data of interest and various information related to data such as quantity, weight and buying frequency is stored. In this paper they has been introduced the concept of mining high utility itemsets considering the distributed environment systems i.e. in case of data of interest is scattered on different locations. For the working on distributed system this paper used master-slave approach. There will be one master site and multiple slave sites. Utilities values of itemsets are extracted on each location. To improve the performance this computation task is done parallel. Then total utility of itemsets is calculated at the master site. They used Fast Utility Itemset algorithm for mining high utility item set. Experiments showed that this reduced the execution time required for the computation.

Vincent S et.al . [23] [27] Frequent Itemset Mining (FIM) is used for discovery of itemsets that customer purchase together frequently. But the model has some problem. It generates high amount of frequent itemsets and avoids low selling frequencies. FIM treats every item with same importance, profit and weight and it assumes an item can be either present or absent i.e. binary representation of itemsets in transaction. To overcome this problem, utility mining concept is emerged. A utility of an itesmset measured in terms of weight, profit, cost, quantity etc. Itemset is considered of a high utility If the utility of an itemset is greater than user-specified minimum utility threshold otherwise it is considered of a low utility. High Utility Mining has several issues like lossy representation of HUIs, extracted pattern reduction may not be achieved, less efficient, recovery of HUIs from concise representation etc. Paper uses closed itemset mining technique and proposed a novel solution viz. AprioriHC, AprioriHC-D, CHUD and DAHU to address these issues. In this paper, they used closed itemset mining technique and proposed a novel solution called Closed High Utility Itemset Mining.

3. CONCLUSIONS

In this paper, we have reviewed various techniques of generating HUIs amongst them Closed High Utility Mining proved better and efficient. But conducting all these processes, organizations have to carry the storage and computation overhead. There is a scope to alleviate the process of generation of compact HUIs to third party and this arise the need of security and privacy of the data as outsourced party may not be trusted one. Further to preserve the privacy we can use the some privacy preservation technique.

4. REFERENCES

[1] R. Agrawal and R. Srikant, “Fast algorithms for mining association rules,” in Proc. 20th Int. Conf. Very Large Data Bases, 1994,pp. 487–499.

[2] Agrawal . R Mannila, H., Srikant R. , Toivonen, H., and Verkamo, A.I. 1996.” Fast discovery of association rules. In Advances in Knowledge Discovery and Data Mining,” AAAI Press, pp. 307–328.

[3]. N. Pasquier, Y. Bastide, R. Taouil, and L. Lakhal, “Efficient mining of association rules using closed itemset lattice,” J. Inf. Syst., vol 24, no. 1, pp. 25–46, 1999.

[4] Fujiwara, S Ullman, J. D. , and Motwani, R.”Dynamic miss-counting algorithms: Finding implication and similarity rules with confidence pruning.” In Proc. of the 16th International Conference on Data Engineering (ICDE’00), San Diego, California, pp. 501–511, 2000.

[5]. Jian Pei,Jiawei Han, “FreeSpan: Frequent pattern projected Sequential pattern mining. School of computing science, Simon Fraser University. 2001.

[6]. T. Calders and B. Goethals, “Mining all non-derivable frequent itemsets,” in Proc. Int. Conf. Eur. Conf. Principles Data Mining Knowl. Discovery, 2002, pp. 74–85.

[7]. J.F. Boulicaut, A. Bykowski, and C. Rigotti, “Free-sets: A condensed representation of Boolean data for the approximation of frequency queries,” Data Mining Knowl. Discovery, vol. 7, no. 1, pp. 5–22, 2003.

[8]. W. Cheung and O.R. Zaı¨ane, “Incremental Mining of Frequent Patterns without Candidate Generation or Support Constraint,” Proc. Seventh Int’l Database Eng. and Applications Symp. (IDEAS ’03), pp. 111-116, 2003.

[9]. J.L. Koh and S.F. Shie, “An Efficient Approach for Maintaining Association Rules Based on Adjusting FP-Tree Structures,” Proc. Ninth Int’l Conf. Database Systems for Advanced Applications (DASFAA ’04), pp. 417-424, 2004

[10] R. C.W. Wong and A.W. Fu. Mining top-k itemsets over a sliding window based on zipfian distribution. In Proc. of SIAM SDM, 2005

[11] .C. Lucchese, S. Orlando, and R. Perego, “Fast and memory efficient mining of frequent closed itemsets,” IEEE Trans. Knowl. Data Eng., vol. 18, no. 1, pp. 21–36, Jan. 2006.

[12]. U. Yun, “Efficient Mining of Weighted Interesting Patterns with a Strong Weight and/or Support Affinity,” Information Sciences, vol. 177, pp. 3477-3499, 2007

[13]. Erwin, A. Gopalan , R. P . Achuthan, N. R.: A Bottom-Up Projection Based Algorithm for Mining High Utility Itemsets. Proceedings of the 2nd International Workshop on Integrating Arti_cial Intelligence and Data Mining, Volume 84, Gold Coast, Australia, 2007.

[14].A. Erwin, R. P. Gopalan, and N. R. Achuthan, “Efficient mining of high utility itemsets from large datasets,” in Proc. Int. Conf. Pacific-Asia Conf. Knowl. Discovery Data Mining, 2008, pp. 554–561.

[15]. K. Chuang, J. Huang, and M. Chen, “Mining top-k frequent patterns in the presence of the memory constraint,” VLDB J., vol. 17, pp. 1321–1344, 2008.

[16].Y.C. Li, J.S. Yeh, and C.C. Chang, “Isolated items discarding strategy for discovering high utility itemsets,” Data Knowl. Eng.,vol. 64, no. 1, pp. 198–217, 2008.

[17]B. Le, H. Nguyen, T. A. Cao, and B. Vo, “A novel algorithm for mining high utility itemsets,” in Proc. 1st Asian Conf. Intell. Inf. Database Syst., 2009, pp. 13–17.

[18]. C. F. Ahmed, S. K. Tanbeer, B.S. Jeong, and Y.K. Lee, “Efficient tree structures for high utility pattern mining in incremental databases,” IEEE Trans. Knowl. Data Eng., vol. 21, no. 12, pp. 1708–1721, Dec. 2009.

[19]. T. Hamrouni, S. Yahia, and E. M. Nguifo, “Sweeping the disjunctive search space towards mining new exact concise representations of frequent itemsets,” Data Knowl. Eng., vol. 68, no. 10,pp. 1091–1111, 2009.

[20].B. Vo, H. Nguyen, T. B. Ho, and B. Le, “Parallel method for mining high utility itemsets from vertically partitioned distributed databases,” in Proc. Int. Conf. Knowl.-Based Intell. Inf. Eng. Syst.2009, pp. 251–260.

[21].V. S. Tseng, C.W. Wu, B.E. Shie, and P. S. Yu, "UP-Growth: An efficient algorithm for high utility itemset mining," in Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining, 2010, pp. 253–262.

[22]. B.E. Shie, H.F. Hsiao, V. S. Tseng, and P. S. Yu, "Mining high utility mobile sequential patterns in mobile commerce environments," in Proc. Int. Conf. Database Syst. Adv. Appl., 2011, vol. 6587, pp. 224–238.

[23] C.W Wu, P. Fournier-Viger, P. S. Yu, and V. S. Tseng, "Efficient mining of a concise and lossless representation of high utility itemsets," in Proc. IEEE Int. Conf. Data Mining, 2011, pp. 824–833.

[24] ..Mengchi Liu, Junfeng Qu, " Mining High Utility Itemsets without candidate generation," School of Computer Science, Carleton University, Canada 2012.

[25]. T. Hamrouni, "Key roles of closed sets and minimal generators in concise representations of frequent patterns," Intell. Data Anal., vol. 16, no. 4, pp. 581–631, 2012

[26]. Kannimuthu Subramanian, " A novel approach to extract high utility itemsets from distributed databases," 2012.

[27]. Vincent S. Tseng, Cheng-Wei Wu, Philippe Fournier-Viger, and Philip S. Yu,"Efficient Algorithms for Mining the Concise and Lossless Representation of High Utility Itemset," Knowl. And data Engg., vol. 27, no.3,2015.