ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 5, Issue 6, June 2016

Privacy Preserving Concise Representation of HUI

Ms.Archana Kisan Dere
Second Year Master of Engineering
Department of Computer Engineering
Sharadchandra Collage Of Engineering
Dumberwadi,Otur,Pune-(India)

Prof. S .A. Kahate
Assistant professor
Department Of computer Engineering.
Sharadchandra Collage Of Engineering.
Dumberwadi,Otur,Pune-(India)

***Abstract*— The market base analysis is the popular application in market. In data mining Frequent Itemset mining is required. More frequent but less important itemset discovered by FIM model having low frequencies. For this problem utility mining is required. Efficiency degraded due to compact size off HUI. To achieve the efficiency closed sets, free set, and compact representation available. The algorithm CHUD is proposed to achieve this. Another party is outsourced to mine the huge organization data. Here is need of preservation of data privacy because third party is not trustworthiness. This task have data privacy problem and this is biggest problem in this system. Proposed system handles the load of computation, storage and processing to another party with preservation of privacy of outsourced high utility mining and also to produce the concise representation of HUIs using existing work. We proposed for privacy preservation one to one substitution method and Homomorphic encryption method for the analysis.**

***Index Terms*— Utility mining, Compact representation of HUIs, Frequent Itemset Mining, HUIs.**

I. Introduction

The popular technique for market basket analysis is FIM (Frequent itemset mining).The Mining of huge amount of data using different technique get the useful and important information. The large amount of frequent itemset generation is the biggest problem is present in FIM model. Less frequent and low importance item generated by itemset will generate high important items of itemsets. FIM treats every item with same importance, profit and weight and it assumes an item can be either present or absent. To overcome this problem, utility mining concept is implemented. 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. The utility of itemset is depends on utility threshold as per requirements.

FIM introduced to reduce cost overhead and mining task and to provide concise representation, different approaches like Freeset, Non derivable sets, Closed Itemset. Applications such as website click, stream analysis, mobile commerce environment, biomedical applications are high utility applications. It is difficult to user for result analysis while using large utility application and it may discover by HUI. It requires more memory for processing. Applying these techniques to HUI produced several challenges:

-Significant reduction in the extracted patterns may not be achieved.
-Lossy representation of all HUI which is not meaningful to user.
-How to recover all HUI's from the concise representation.

-Algorithms may not be efficient.

An algorithm discussed integration of closed utility itemset into high utility itemset is proposed which is also known as closed +high utility itemset. Storage of all itemsets and utility information requires extra memory and extra processing task to get a set of CHUI this is the highly computational task. The huge amount of data is tends to outsourced to the outsourced party. In existing systems, the outsourced party will have the information about High Utility Items and their so does not preserve the data privacy and is the biggest problem. Proposed system is to handle the load of computation, storage and processing to another property with preservation of privacy of data.

II. Literature Survey

R.agarwal et.al.[1] In database mining, association rules discovery is an important problem. This paper 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 . [2] 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.

Fujiwara et.al.[3] 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.

T. Calders et.al [4] 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 [5] 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.

Y.C. Li et.al.[6] From business perspective, itemset share values reflect more the significance of itemsets for mining association rules in a database. The Share-counted FSM (ShFSM) algorithm is one of the best algorithms which can discover all share-frequent itemsets efficiently. But ShFSM wastes the computation time on the join and the prune steps of candidate generation in each pass, and generates too many useless candidates. Therefore, this study proposes the Direct Candidates Generation (DCG) algorithm to directly generate candidates without the prune and the join steps in each pass. Moreover, the number of candidates generated by DCG is less than that by ShFSM. Experimental results reveal that the proposed method performs significantly better than ShFSM.

C. Lucchese et.al .[7] Proposed High Utility Pattern mining had been several applications in broader aspect. Efficient mining of high utility itemsets was very important in data mining. Paper 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.
C. F. Ahmed et.al. [8] 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. [9] was introduced exact concise representation based on discovery of the disjunctive search space. The disjunctive 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.
V. S. Tseng et.al.[10] 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. This paper 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.[11] 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

Mengchi Liu et.al. [12] 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.

Kannimuthu Subramanian et.al.[13] 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. Paper used Fast Utility Itemset algorithm for mining high utility item set. Experiments showed that this reduced the execution time required for the computation.

Zheng et.al.[14] , integrated the concept of utility into sequential pattern mining, and defined the generic framework for high utility sequence mining. This paper presented an efficient algorithm called USpan, which is used to mine high utility sequential patterns. USpan used lexicographic quantitative sequence tree to extract the complete set of high utility sequences and designed concatenation mechanisms for calculating the utility of a node and its children with two effective pruning strategies. Experiments on both synthetic and real datasets showed that USpan efficiently identifies high utility sequences from large scale data with very low minimum utility.

. Vincent S et.al [15]. 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. 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.

For discovery of itemsets Frequent Itemset Mining is used that customer purchase together frequently. But the model has some problem. It generates high amount of frequent itemsets and avoids low selling frequencies. To overcome this problem, utility mining concept is emerged. Paper [10] used closed itemset mining technique and proposed a novel solution called Closed High Utility Itemset Misning. Methodology in this paper is extends further to achieve privacy in the outsourced mining task. 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.

III. System Architecture

A. Existing System

Extracting less and most important high utility item sets and used various strategies to enhance the performance of the mining task. The algorithm Closed High Utility Item sets Discovery includes:

1. Procedure

A - Converts the database in vertical database and simultaneously calculates the utility for each transaction and also its transaction weighted utility of items. When transaction is fetched, its Tid and transaction utility are stored into global table named Global Utility Table (GUT).

B – In this algorithm process to scans database and collect promising items having estimated utility greater than abs_min_utility into ordered list which is sorted in increasing order of support. Then utilities of unpromising items are removed from Global Utility table (GUT).

C – In this step CHUD generates candidates in recursive manner starting from candidates containing a single promising item and recursively joining items to them to form larger candidates. Using the total order α complete set of itemsets can be divided into n non-overlapping subspaces, where the kth subspace is the set of itemsets containing the item ak but no item ai α ak. For each item ak ε O, CHUD creates a node N({ak}) and puts items a1 to ak-1 into PREV-SET({ak}) and items ak+1 to an into POST-SET({ak}). Then CHUD calls the CHUD Phase-I procedure for each node N({ak}) to produce all the candidates containing the item ak but no item ai α ak.

Phase I

A - Here performs Subsume check on X which verifies if there exists an item which is included in a closed item set that has already been found and supersets of X do not need to be implemented.

B – In this to compute the closure of an Xc = C(X) of X. Then the estimated utility is calculated.

C – After that DCM strategy (Discarding candidates with maximum utility less than minimum utility threshold) is applied i.e. it computes the maximum utility of Xc. It discards the candidate whose estimated or maximum utility is less abs_min_utility; otherwise Xc is outputted with its estimated utility.

D – Here a node N (Xc) is created and the procedure Explore is called for finding candidates that are supersets of Xc i.e. potential CHUIs. Then RML strategy is applied to remove minimum utility items from local transaction utility table.

Phase II

A - Consists of taking each candidate X and calculate its utility and generate the utility unit array. Each candidate of low utility is discarded.

B. Proposed System

Mining concise representation of HUIs, methodology in [10] is used. Privacy in outsourced task is achieved by proposed system. Proposed techniques are applied before the mining task is outsourced to another party. This preprocessing contains following methods:

- One to one substitution for items in the itemsets.
- Homomorphic encryption of weights of the data.
- Closed High utility item set mining on transaction data using encrypted utility threshold and Homomorphic encryption properties.

A. one to one substitution: - every item in the transaction table is substituted with its respective numerical Id. Each item in item set is assigned with numerical id randomly.

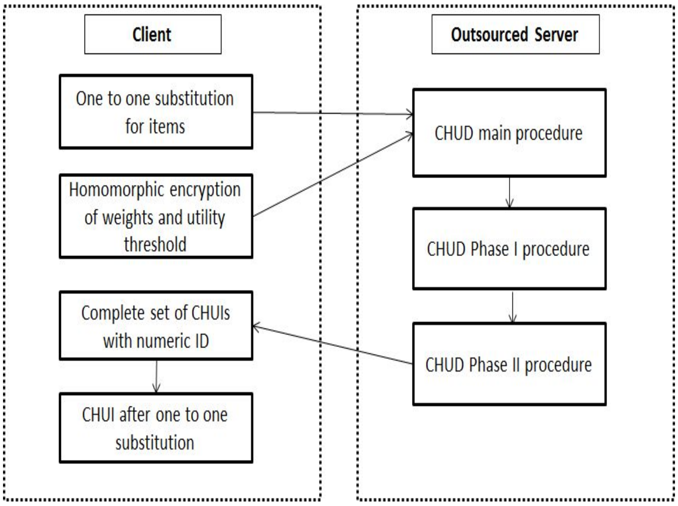


Fig. 1. Proposed System

Output of this process is converted dataset D', Map of the items and their numerical ids. This map is kept securely at client side and converted dataset D' is forwarded to next step. Example of this step is given below.

Table 1: One to One substitution

D	D'
I1,I2,I3	78,421,568
I4,I1	142,78
I5,I2,I1,I3	2,421,78,568
I1,I2	78,421
I2,I4	421,142

Consider data set D shown in table 1, I is item set which contains {I1, I2, I3, I4, I5} and D has 5 transactions as shown in first column of Table 1. Each Item I1 to I5 is assigned one numeric ID randomly and D is converted D' as shown in Second column of table 1 and Map M of this assignment is stores locally at client

Table 2: Map M of the Item and Its numeric ID

Item	Random Numeric ID
I1	78
I2	421
I3	568
I4	142
I5	2

B. Homomorphic encryption of weights of the data:- In proposed scheme, weights of the items are encrypted using homomorphic cryptosystem. Proposed system will use homomorphic encryption because it allows mathematical operation on encrypted data and when results of this mathematical encryption are decrypted then decrypted results reflects the mathematical operation's effect.

Here weights of the items are the important information in High utility item set mining; in existing system weights of the items are understandable to outsourced party therefore indirectly HUI are understandable to the outsourced party.

Proposed System use one of the Pallier cryptosystem homomorphic cryptosystem, which exhibits following properties:

i. Homomophic addition

Dsk (Epk (a+b)) = Dsk(Epk(a) * Epk(b) mod N^2)

ii. Homomorphic Mulitplication

Dsk(Epk (a*b)) = Dsk(Epk $(a)^b$ Mod N^2)

Where Epk is encryption function with Key public key Pk derived using N and g where N is product of two prime numbers of similar two lengths and g is generator in Z_{N2}. Also, let Dsk be the decryption function with secret key sk. Encrypted weights of the items, Converted Database and encrypted utility threshold given to third party for high utility mining.

iii. Semantic Security: It is impossible to figure out information about plain text using cipher text.

For mining the database is given to outsourced third party.

C. CHUI Mining: -

This outputs the complete set of CHUIs which is in substituted form, and then at client side set of CHUIs is obtained using locally kept map of items and its numerical Ids. We use mining of CHUIs as per the methodology described in the section III.

.

IV. Mathematical Model of Proposed System

S be a Privacy Preserving Concise High Utility Itemset Representation system having Input, Processes and Output. It can be represented as,

S = {I, P, O}

Where, I is a set of all inputs given to the System, O is a set of all outputs given by the System, P is a set of all processes in the System.

I = {I1, I2, I3, I4, I5, I6, I7}

I1 - Database containing Itemsets and Transactions.

I2 - Minimum Utility Threshold and item weight

I3 - Input D', encrypted weight and utility threshold.

I4 - Items and its estimated utility.

I5 - Ordered List of promising items.

I6 - Node N(X), GTU table and abs_min_utlity.

I7 - Set of potential CHUIs.

I8 – Set of CHUIs.

P = {P1, P2, P3, P4, P5, P6, P7, P8}

P1 - Convert the database D by one to one substitution method i.e. substitute items with random numeric Ids and also store the map of item sets and its numeric id locally.

P2 - Encryption of weight using Homomorphic encryption and generate encrypted utility threshold.

P3 - CHUD main procedure scans the database and converts into vertical database and at the same time computes the transaction utility for each transaction and calculates TWU of items. Transaction utility and its Tid are loaded into Global TU-table.

P4 - Items having estimated utility greater than abs_min_utility are called promising items and collected in ordered list O sorted according to increasing order of support and then utilities of unpromising items are removed from the GTU.

P5 – CHUD generates candidates containing single item and recursively joining items to them to form larger candidates. It creates Node N({ak}) with PREV-SET({ak}) and POST-SET({ak}).

P6 – CHUD Phase I procedure called for each node. It computes Closure of an item and explores the tree until all pCHUIs generated.

P7 – CHUD phase II takes each itemset X and calculates its utility and utility unit array. Low utility candidate is discarded.

P8 - The complete set of CHUIs which is in substituted form, is obtained using locally (at client side) kept map of items and its numerical Ids.

O = {O1, O2, O3, O4, O5, O6, O7,O8}

O1 - Converted database D'

O2 - Encrypted weight and utility threshold.

O3 - Global TU-table

O4 – Ordered list (O) of promising items

O5 – Node N({ak})

O6 – Set of pCHUIs

O7 – Set of CHUIs

O8 – Obtained set of CHUIs from locally stored map.

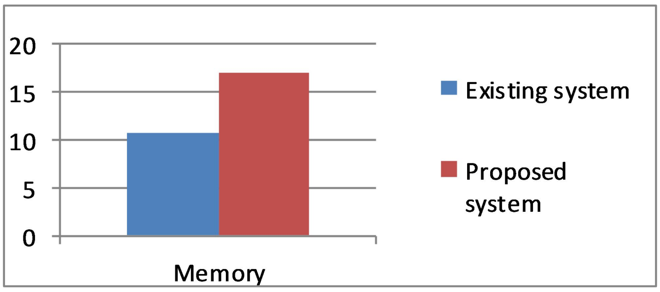
V. Experimental Setup and Results

Experiment showed the effect of applying privacy preservation methods and also showed the time and memory requirements in existing CHUI generation method and the proposed Secure CHUI generation method.

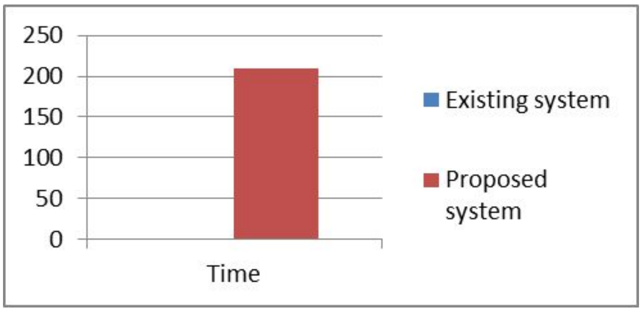
The existing work takes 0.46 sec time and 10.77 MB memory to complete a task and proposed method takes 209 sec time and 17 MB memory respectively. The time and memory requirement for proposed system is more than existing method. This is because of application of privacy preservation techniques on available transaction database before CHUI generation. For experiment purpose, Footmark dataset is used for both approaches containing 100 transactions each. It's a real life dataset, each transaction contains up to 8 items. The minimum utility threshold is randomly given from 1000 to 2000. Table shows the expected results. Graph 1, 2 and 3 shows the graphical representation of the results.

Table 3: Comparison between existing and proposed work.

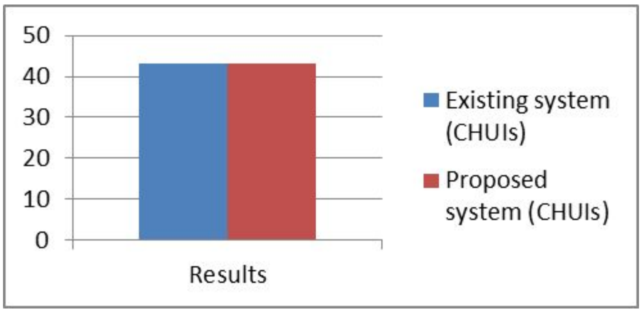
Parameters	Proposed method (min.utility threshold -1000)	Existing method (min.utility threshold -1000)
Result	43 (CHUIs)	43 (CHUIs)
Time (sec)	209 sec	0.46
Memory (MB)	17 MB	10.77 MB
Privacy preserved?	YES	NO



Graph 1: Comparison of memory requirement between existing and proposed system.



Graph 2: Comparison of time requirement between existing and proposed system.



Graph 3: Number of CHUIs generated for existing and proposed system

VI. CONCLUSION

From the experimental result we conclude that even though existing system was efficient than our proposed system in case of time and memory but it was not secure if we want to send data to outsource party for doing computation and storage of HUI's. Our proposed system doing process of generation of compact HUIs to another party and also fulfilled the aroused need of security and privacy of the data as outsourced party may not be trusted one.

VII. FEATURE SCOPE

In our system CHUI's discover with respect to Minimum Threshold Utility. So in future we can implement our system which will select the optimu Minimum utility threshold to avoid the problem of this system such as discover more number of item set's when giving low value for Minimum utility threshold vice versa.

REFERENCES

[1] 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.

[2] 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

[3] Fujiwara, S., Ullman, J.D., and Motwani, R. 2000. "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.

[4] 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.

[5] 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.

[6] Y.C. Li, J.S. Yeh, and C.-C. Chang. "Direct candidates generation: A novel algorithm for discovering complete share-frequent itemsets." In Proc. Fuzzy Systems and Knowledge Discovery, pages 551–560, 2005.

[7] 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. 2008.

[8] 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.

[9] Engelbert Mephu Nguifo, Sadok Ben Ben Yahia "Sweeping the Disjunctive search space towards mining new concise exact representations of frequent itemstes," article in data & knowledge engineering · 2009.

[10] V. S. Tseng, C.-W. Wu, B.E. Shie, and P. S. Yu. "Upgrowth: An efficient algorithm for high utility itemset mining." In Proc. ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining, pages 253–262, 2010.

[11] 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.

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

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

[14] Junfu Yin, Zhigang Zheng, "USpan: An Efficient Algorithm for Mining High Utility Sequential Patterns," KDD'12, August 12–16, 2012, Beijing, China.

[15] 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.

Archana Kisan Dere I have completed Diploma in Computer Engineering in 2008 from Institute Of Technology,kuran in Mumbai. University. I have also completed Bachelor Of Computer Engineering in 2011 from D. N. Patel College of Engineering , Shahada in North Maharashtra University.

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 5, Issue 6, June 2016

Prof. Sandip A. Kahate
B.E. in computer science and engineering frc Amravati university, M.E. in Wirele Communication and Computing, from Nagp University and preparation for Ph. D. registration. I is currently working as an Assistant Professor Computer Engineering Department, Sharadchanc Pawar College of Engineering, At. Post-Ot Dist-Pune-412409(M.S.), India. He has 10 years teaching experience. He is author of 1 research pap with around 10 papers in international journal and 3 international conference in India and abroad. His are of interest are Wireless communication a computing, network security and Ad-Hoc Network.