IJETST- Vol.||04||Issue||09||Pages 6065-6066||September||ISSN 2348-9480 2017

International Journal of Emerging Trends in Science and Technology
IC Value: 76.89 (Index Copernicus) Impact Factor: 4.219 DOI: https://dx.doi.org/10.18535/ijetst/v4i9.32

# Efficient High Utility Top-K Frequent Pattern Mining

Authors
**J. Krishna[1], M. Rupesh Kumar Reddy[2], Dr. M. Rudra Kumar[3]**
[1]Assistant Professor, AITS, CSE Research Scholar, RU, Kurnool
[2]M.Tech.,(PG Scholar), Dept of CSE,Annamacharya Institute of Technology & Sciences, Rajampet, Kadapa,
[3]Professor, Dept of CSE, Annamacharya Institute of Technology & Sciences, Rajampet, Kadapa

**Abstract**
High utility pattern mining can be defined as discovering sets of patterns that not only co-occurs but they carry high profit. In two-phase pattern mining an apriori algorithm is used for candidate generation. However candidate generation is costly and it is challenging problem that if number of candidate are huge then scalability and efficiency are bottleneck problems. As a rule, finding a fitting least utility edge by experimentation is a monotonous procedure for clients. In the event that min_util is set too low, an excessive number of HUIs will be produced, which may bring about the mining procedure to be exceptionally wasteful. Then again, if min_util is set too high, it is likely that no HUIs will be found. In this paper, we address the above issues by proposing another structure for top-k high utility thing set mining, where k is the coveted number of HUIs to be mined. Two sorts of proficient calculations named TKU (mining Top-K Utility thing sets) and TKO (mining Top-K utility thing sets in one stage) are proposed for mining such thing sets without the need to set min_util. We give an auxiliary examination of the two calculations with talks on their preferences and restrictions. Exact assessments on both genuine and manufactured datasets demonstrate that the execution of the proposed calculations is near that of the ideal instance of best in class utility mining calculations.
**Keywords:** Utility mining, high utility item set mining, top-k pattern mining, top-k high utility item set mining, Data mining, frequent itemset, transactional database.

## 1.Introduction

Discovering information of products in the market is a crucial task which required complex analysis of information of each product from the user perspective. Identify frequent itemsets from huge database required valuable study. Apriori algorithm is most popular algorithm for pattern mining. It is breadth first search algorithm which scan database as many times as the length of frequent pattern. Author in [2] discussed about the strategy of FP-mining in which frequent pattern tree without generating any candidate gives the frequent itemsets without any candidate key and searches the database only twice. J. Han discussed about FP-pattern tree. It is squeeze form of extended prefix-tree to extract crucial information about frequent patterns. It is depthfirst algorithm. FP-growth algorithm used partitioned based method for decomposing the mining task into small set of task [7]. Weighted association ruled framework mines the frequent itemsets by considering its weight. It is outstanding framework but does not considered importance of item quantities in transaction DB [12]. In high beneficial itemset mining considered items frequency, weight and efficiency. It has issue that it generates the large number of candidate which is infancy task. Mining plays an important role to extract hidden information from large dataset. Transaction weighted utilization model is discussed in [2], is used to overcome the problem in pruning search high utility itemsets. It is difficult task as there may chances that superset of low utility can be high utility. Two phase candidate generation algorithm is implemented for utility itemset mining [3] [6] [9] [11] in which first phase extract the candidates of high utility patterns and then again scan the unrefined data for finding more candidates. It is the problematic issue as number of candidates get increased which may caused efficiency and scalability issue. Therefore it degrades the efficiency of system. For mining high utility itemsets HUIMiner algorithm is utilized in [10]. This algorithm is less efficient when there is requirement huge database for mining due its inefficiency of join operation, pruning approach as well as scalability issues with the vertical dataset structure. D2HUP, algorithm is finding to be novel solution for mining utility itemsets in share framework. This algorithm can addresses the scalability and efficiency issues occurred in the existing systems as it directly extracts the high utility patterns from large transactional databases i.e. TWU. Strength of D2 HUP algorithms is based on the powerful pruning approaches. It tries to find the patterns in recursive enumeration and it utilizes the singleton and closure property to enhance the efficiency of dense data. Linear data structure known as CAUL is used to show the original information of utility in the unrefined data, it helps to discover the root causes of prior algorithm which employs to maintain data structure information of original utility. Constraints based mining is derived approach from frequent pattern mining to mining utility. Its major is to push the constraints into the frequent pattern mining. In [14], constraints are defined same like normalized weighted support. To prune the search space DualMiner algorithm is introduced in [4] with antimonotone & monotone constraints. In [5], author De Raedt et al. defined the way of applying standard constraint programming techniques on constraint based mining issues. There are some categories of utility mining defined in [17] such as, objective, subjective and semantic measures. Objective measures can be defined as confidence or support for data. In subjective measure unexpectedness is considered which take knowledge domain of users account and the semantic measures are also called as, utilities which consider the both data and user expectations. Among all these three categories are discussed in [8][9] and [13]. In this first objective measure considers the product price in shopping basket then subjective measures are based on count and amount shares. The utility measure is equivalent to the both objective and subjective measures. The concept of weighted itemset mining and association rule mining is proposed by Cai et al and Lin et al. in [4]. A vertical weight is the concept of significance transaction discussed in [12].

High utility mining is the new approach for the problem which mainly concentrates on high utility pattern growth mechanism. In this approach to reduced the drastically pattern growth reverse set enumeration approach and pruning algorithm is proposed. But it is practically impossible to enumerate all patterns and prune search space. Therefore, anti-monoticity property is applied to pruning approach. Pattern growth approach estimates upper bound of utilities of possible patterns represented by nodes of rooted subtree. Frequent itemset mining means finding items that presents in a database above a user given frequency threshold value. These techniques do not consider the quantity of items or profit of the purchased items. Accordingly it is not efficient for the end-user who want find the importance of the items in database. After all quantity of items and profit of items are basic terms for maximizing the profit of the organization. For this purpose new technique in data mining is introduced called as high utility mining. This technique is useful for finding itemsets from database which gives high utility. Utility means influence or usefulness of items. Utility of items is mainly calculated by multiplying internal utility and external utility. Itemset in a single transaction is called utility or internal utility and itemset in different transaction database is called external utility.It allows users to identify the usefulness or importance of items using different values. Thus, it emulates the impact of different items. High utility itemsets mining is useful in decision making activity of many applications, such as retail marketing and Web service, since items are actually different in many ways in real applications. High utility itemset is itemset which having utility no less than a user-specified minimum utility threshold value; otherwise, it is called a low-utility itemset. In many applications such as cross-marketing in retail stores mining such high utility itemsets from databases is an important process. Existing techniques which [2, 3, 4, 5, 6, 7] used for utility pattern mining in large database. However, the existing methods usually generate a large set of potential high utility itemsets and the mining performance is degraded consequently.

## II Literature Survey

In the literature survey we will go to discuss various existing methods which allow user to access the services from multiple service providers in High

Utility Itemsets Mining. Below we are discussing some of them. Chowdhury Farhan Ahmed, Syed Khairuzzaman Tanbeer, Byeong-Soo Jeong, and Young-Koo Lee granted three novel tree structures for efficiently perform transactional and interactive HUP mining [2]. The first tree structure is used to organize the items according to their lexicographic order. It admitted as Transactional HUP Lexicographic Tree (IHUPLTree). It captures the transactional data without any restructuring operation. The next tree structure is the IHUP Transaction Frequency Tree (IHUPTF-Tree), which is useful arranging items according to their transaction frequency in descending order. To curtail the mining time, the last tree, IHUP-Transaction-Weighted Utilization Tree (IHUPTWUTree) is designed. The structure of this tree is based on the Transactional Weighted Utility(TWU) value of items in descending order. Alva Erwin, Raj P. Gopalan, and N. R. Achuthan, Advised CTU-PROL algorithm for efficient mining of high utility itemsets from large datasets[3]. These algorithms search the large TWU items in the transaction database. If data sets is too large to be held in main memory, the algorithm generates subdivisions using parallel projections and for each subdivision, a Compressed Utility Pattern Tree (CUP-Tree) is used to mine the complete set of high utility itemsets. If the dataset is Limited, it built a single CUP-Tree for mining high utility itemsets. Shankar S., Purusothaman T., Jayanthi, S., proposed a novel algorithm for mining high utility itemsets[4]. This fast utility mining (FUM) algorithm finds all high utility itemsets within the disposed utility constraint threshold. The proposed FUM algorithm scales strong as the capacity of the transaction database increases with regard to the number of distinct items available. R. Chan, Q. Yang, and Y. Shen, implied mining high utility itemsets[5]. They proposed a novel concept of top-K objective directed data mining, which spotlights the top-K high utility closed patterns. They compute the concept of utility to capture highly desirable statistical patterns and present a level wise itemset mining algorithm. They create a new sniping strategy based on utilities that allow pruning of low utility itemsets to be done by means of a anemicer but antimonotonic condition. Ramaraju C., Savarimuthu N., implied a conditional tree based novel algorithm for high utility itemset mining [6]. A novel conditional high utility tree (CHUT) reduce the transactional databases in two stages to compress search space and a new algorithm known as HU-Mine is proposed to mine complete set of high utility item sets. Y. Liu, W. Liao, and A. Choudhary, implied a fast high utility itemsets mining algorithm [7]. They are proposed a Two-Phase algorithm to efficiently snip down the number of candidates and can precisely obtain the complete set of high utility itemsets. In the first aspect, they propose a model that applies the transaction-weighted downward closure property on the search space to facilitate the identification of candidates. Recent phase identifies the high utility itemsets. Adinarayanareddy B., O. Srinivasa Rao, MHM Krishna Prasad, implied improved UP-Growth high utility itemset mining[8]. The compact tree structure, Utility Pattern Tree i.e. UP-Tree, maintains the history of transactions and their itemsets. It expedites the mining performance and avert scanning original database frequently. UP-Tree scans database only two times to achieve candidate items and manage them in an efficient data structured way. UP-Growth gates more execution time for Second Phase by prating UP-Tree. Hence they proposed modified algorithm aiming to reduce the execution time by effectively identifying high utility itemsets. P. Asha, Dr. T. Jebarajan, G. Saranya, implied a survey on efficient transactional algorithm for mining high utility itemsets in distributed and dynamic database [9]. The proposed system applies one master node and two slave nodes. Database is subdivided for every slave node for computation. The slave node measures the existence of each item. These datas are gathered in their local table. Then each slave node forwards these tables to master node. The Master Node maintain global table for gathering these data. Based on the minimum utility threshold value it measures the promising and unpromising itemsets. Ahmed et al. [10] implied a structure named IHUP-Tree for maintaining essential information about utility mining. It avoids scanning of database for multiple times and generating candidates or patterns during the mining process. However, although IHUP-Tree produces better performance than Two-Phase and IIDS, it still provides too many HTWUIs. Tseng et al. proposed a novel algorithm named UP-Growth [11], which applies several pruning and counting strategies during the data mining processes. By the proposed strategies, the estimated utilities are effectively decreased in UP-Trees during the data mining processes and the number of HTWUIs is further reduced. Therefore, the system performance of utility mining can be improved significantly. In

1994, R.Agrawal and R. srikant [1], discussed about the problems of extracting association rules between the items in huge databases of sales transactions. Two algorithms namely, Apriori and AprioriTid are proposed in this paper to solve the problem with other algorithms. Both algorithms are integrated together for hybrid algorithm. It is known as, "AprioriHybrid" algorithm. AprioriHybrid algorithm has its own scalability properties. Another is the problem of basket data is also discussed in this paper. It contains the huge applications database. To make discovery of n-number of itemsets there is need of multiple passes over the data. At the very beginning, it determines the individual itemset which has minimum support. The proposed algorithm Apriori and AprioriTid are different from the AIS and SETM algorithms with respect to candidate itemsets. In AprioriTid algorithm one additional property is used to count the support of candidate itemsets after initial pass. For three datasets performance of AprioriHybrid is relevant to the Apriori and AprioriTid algorithm. In all cases the proposed AprioriHybrid outputs the better performance rather than the Apriori. In the last pass switches AprioriHybrid performs the little worst than the Apriori algorithm. Therefore, AprioriTid algorithm is used after each space. In 1998,R. Hilderman, Colin L. et al. [2], proposed shared confidence framework. It is the framework to discover the knowledge from databases. It also addresses the problem of discovering itemsets from market basket data. In this paper, they concerned on two types of goals such as, first one is to introduce measures of itemset. This measure is useful and practically interactive for commonly used support measure. Secondly, the discovery of profiles of customers buying patterns also to discover profiles of customer which is done by splitting them into individual classes. The proposed mechanism merged the Apriori algorithm to make discovery of association rules between large databases itemsets. In this paper, an experimental results analysis represented that the proposed share confidence framework has ability to give more information feedback than the support confidence framework. In 2000, M.J. Zaki, C.J. Hsiao [3], represented CHARM. It is an efficient algorithm for mining closest frequent itemsets. The frequent pattern mining includes the discovery of association rules, powerful rules, multidimensional patterns and also other important discovery. To addressed the problem in frequent pattern mining. An apriori algorithm is employs the BFS i.e. Breadth First Search to enumerates the individual frequent itemsets. Downward closure property is used by apriori algorithm to prune the search space. For mining long patterns there two type of solutions are given in this paper, from those solution first is to discover maximum frequent patterns which has the fewer magnitude than all frequent patterns whereas, the other solution mines frequent closed itemsets. The proposed algorithm CHARM, discovered the itemsets and transaction space over novel tree called as, itemsettidset tree (IT). It uses hash-based approach to eliminate non-closed itemsets at the time of subsumption checking. The algorithm is introduced in this paper is CHARM-L to construct a structure of itemsets. It utilizes the intersection-based approach to non-closed itemsets at the time of subsumption checking. For consideration of appeared IT pairings in the prefix class CHARM-EXTEND is responsible. CHARM-EXTEND mainly return the set of closed frequent itemset. In 2004, J. Pei, J. Han et al [4], discussed about FPgrowth algorithm. In this paper, they mainly contribute themselves to show appropriate order of items. In this paper, author represented the effectiveness of the proposed algorithm. The proposed algorithm is systematic way to incorporate two stages of classes' constraints. In this paper, the concept of convertible constraints is introduced. The convertible constraints are divided into three classes such as, convertible antimonotone, convertible monotone and strongly convertible. Using this number of useful constraints is covered. The convertible constraints cannot be pushed into fundamental apriori framework but they can push into frequent pattern growth mining. Therefore, they were developed fast mining algorithm for various constraints for mining frequent pattern. In 2005, Ying Liu, W.K. Liao [5], represented the ARM i.e. Association Rule Mining technique. It discovers the frequent itemsets from the large database and considered individual item to generate association rules. ARM only reflects impact of frequency of the presence and absence of an item. An anti-monotone property is used to discover frequent itemsets. Mining using Expected Utility (MEU) is used to prune the search space by anticipating the high utility k-itemsets. In the section of experimental analysis they analyzed the scalability and accuracy of results. Finally it is seems that in this paper, Two-phase algorithm can efficiently extract HUI. In 2006, L. Geng, H. Hamilton [6], studied the frequent itemsets. They proposed a best well known algorithm for discovering frequent itemsets. Apriori algorithm is used for pruning search space of itemsets. In this paper, different interestingness is measures of domain of data mining have been proposed. There are three objectives discussed in above from them subjective and semantic based measures deals with background knowledge and goals of user's. These measures are suitable for user experience and the interactive data mining. But the problem in the area of frequent mining is that the real human interest remains an open challenging issue. The experimental setup shows that the human needs to measure their interestingness using another method of analysis. User interactions are crucial in the identification of rule interestingness. In 2008, A. Erwin, R.P. Gopalan et al. [7], proposed TWU algorithm. This algorithm is based on compact utility pattern tree data structures. It implements the parallel projection scheme to utilized disk storage. The algorithm CTU-Mine is proposed for mining HUI from the huge datasets. This algorithm first identifies the TWU items from transaction database. CUP-Tree is the Compressed Utility Pattern Tree for mining complete set of high utility patterns. This algorithm used parallel projection to create subdivision for subsequently mining. TWU has anti-monotone property which is used to discover the pruning space. In this paper the task of HUI mining discovers all the utility which has utility higher than the user specified-utility. CTU-PROL works against the Two Phase algorithm as well as CT Mine. Efficiency of CT-PROL algorithm is improved than the CTU-Mine. In future work to reduce the computation in large database mining they planned to implement a sampling based approximation. In 2008, Yu-Chiang Li, Jieh-Shan Yeh [8], proposed IIDS i.e. Isolated Items Discarding Strategy. It is implemented to address the problem in previously proposed apriori pruning algorithm which cannot identify high utility itemsets. The proposed IIDS is utility mining algorithm; it reduces the candidates and enhanced the performance. In this paper, IIDS to ShFSM and DCG applies two methods FUM and DCG+. These methods are implemented respectively. IIDS provides an efficient way to designed critical operations by using transaction weighted downward closure. The proposed IIDS can be applied on traditional Apriori algorithms to extend the scope of IIDS to specific classification model. In further implementation they discussed about classification problems in data mining. They were planned to combined classification and the association rule mining i.e. established the connection between mining utility and associative classification. In 2011, A. Silberschatz, A. Tuzhilin and T.D.Bie [9], classified the measure into actionable, unexpected and examined the relationship between them. They represented the MaxEnt model. It is used to swap randomization and hence it is computationally more efficient. In this paper, a MaxEnt model is proposed for efficient computations. In this paper, they outlined different ways in the MaxEnt model that can be used efficiently for sampling random databases which is helpful to satisfy the prior information. The parallel to this work, in this paper author made the investigation of MaxEnt modeling strategy for different types of data like, relational databases. In 2016, Junqiang Liu, Ke Wang, Benjamin et al. [10], suggested D2HUP, algorithm. It seems to be novel solution for mining utility itemsets in share framework. This algorithm can addresses the scalability and efficiency issues occurred in the existing systems as it directly extracts the high utility patterns from large transactional databases i.e. TWU. Strength of D2 HUP algorithms is based on the powerful pruning approaches. It tries to find the patterns in recursive enumeration and it utilizes the singleton and closure property to enhance the efficiency of dense data. Linear data structure known as CAUL is used to show the original information of utility in the unrefined data, it helps to discover the root causes of prior algorithm which employs to maintain data structure information of original utility. Constraints based mining is derived approach from frequent pattern mining to mining utility. Its major is to push the constraints into the frequent pattern mining.

## III. Proposed Work

Frequent item set mining is a fundamental research topic in data mining (FIM) mining. However, the traditional FIM may discover a large amount of frequent but low-value item sets and lose the information on valuable item sets having low selling frequencies. Hence, it cannot satisfy the requirement of users who desire to discover item sets with high utilities such as high profits. To address these issues, utility mining emerges as an important topic in data mining and has received extensive attention in recent years. In utility mining, each item is associated with a utility (e.g. unit profit) and an occurrence count in each transaction (e.g. quantity). The utility of an item set represents its importance, which can be measured in terms of weight, value, quantity or other information depending on the user specification. An item set is called high utility item set (HUI) if its utility is no less than a user-specified minimum utility threshold min_util. HUI mining is essential to many applications such as streaming analysis, market analysis, mobile computing and biomedicine

1. Efficiently mining HUIs in databases is not an easy task because the downward closure property used in FIM does not hold for the utility of item sets.
2. In other words, pruning search space for HUI mining is difficult because a superset of a low utility item set can be high utility.

The concept of transaction weighted utilization (TWU) model was introduced to facilitate the performance of the mining task. In this model, an item set is called high transaction-weighted utilization item set (HTWUI) if its TWU is no less than min_util, where the TWU of an item set represents an upper bound on its utility. Therefore, a HUI must be a HTWUI and all the HUIs must be included in the complete set of HTWUIs. A classical TWU model-based algorithm consists of two phases. In the first phase, called phase I, the complete set of HTWUIs are found. In the second phase, called phase II, all HUIs are obtained by calculating the exact utilities of HTWUIs with one database scan.

Advantages Of Proposed System:

1. Two efficient algorithms named TKU (mining Top-K Utility items ets) and TKO (mining Top-K utility item sets in one phase) are proposed for mining the complete set of top-k HUIs in databases without the need to specify the min_util threshold.
2. The construction of the UP-Tree and prune more unpromising items in transactions, the number of nodes maintained in memory could be reduced and the mining algorithm could achieve better performance.

## V **Methodology**

Figure 1 represents the system architecture. In this there are three entities presented such as, user, transaction, HUIMiner. 1. User: -User uploads the transaction dataset. -Get HUI itemsets
2. Transaction:
-Save user uploaded transaction dataset.
-Generate XUT
-Generate reverse set enumeration tree
- Travel tree using DFS
-Search patterns
-Get d2HUP value.
-Verify threshold
-Add itemset in HUI 1. HUIMiner:
- Calculate item relevance score
-Calculate upper bound
-Apply pseudo random projection -Send d2HUP value to transaction entity.

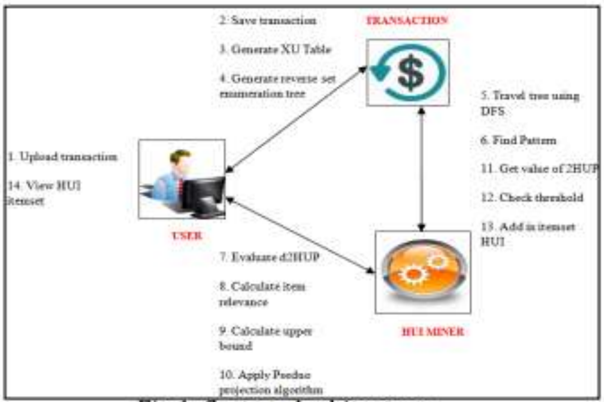

Fig 1: System Architecture

## VI. CONCLUSION

In this paper, we have studied the problem of top-k high utility item sets mining, where k is the desired number of high utility item sets to be mined. Two efficient algorithms TKU (mining Top-K Utility item sets) and TKO (mining Top-K utility item sets in One phase) are proposed for mining such item sets without setting minimum utility thresholds. TKU is the first two-phase algorithm for miningTop-k high utility item sets, which incorporates five strategies PE, NU, MD, MC and SE to effectively raise the border minimum utility thresholds and further prune the search space. On the other hand, TKO is the first one-phase algorithm developed for top-k HUI mining, which integrates the novel strategies RUC, RUZ and EPB to greatly improve its performance. Empirical evaluations on different types of real and synthetic datasets show that the proposed algorithms have good scalability on large datasets and the performance of the proposed algorithms is close to the optimal case of the state-of-threat two-phase and onephase utility mining algorithms.

## Reference


1. R. Agrawal and R. Srikant, “Fast algorithms for mining associationrules,” in Proc. Int. Conf. Very Large Data Bases, 1994, pp. 487–499.
2. C. Ahmed, S. Tanbeer, B. Jeong, and Y. Lee, “Efficient tree structuresfor high-utility pattern mining in incremental databases,”IEEE Trans. Knowl. Data Eng., vol. 21, no. 12, pp. 1708–1721, Dec.2009.
3. K. Chuang, J. Huang, and M. Chen, “Mining top-k frequent patternsin the presence of the memory constraint,” VLDB J., vol. 17,pp. 1321–1344, 2008.
4. R. Chan, Q. Yang, and Y. Shen, “Mining high-utility item sets,” inProc. IEEE Int. Conf. Data Mining, 2003, pp. 19–26.
5. P. Fournier-Viger and V. S. Tseng, “Mining top-k sequentialrules,” in Proc. Int. Conf. Adv. Data Mining Appl., 2011, pp. 180–194.
6. P. Fournier-Viger, C.Wu, and V. S. Tseng, “Mining top-k associationrules,” in Proc. Int. Conf. Can. Conf. Adv. Artif. Intell., 2012, pp. 61–73.
7. P. Fournier-Viger, C. Wu, and V. S. Tseng, “Novel concise representationsof high utility item sets using generator patterns,” inProc. Int. Conf. Adv. Data Mining Appl. Lecture Notes Comput. Sci.,2014, vol. 8933, pp. 30–43.
8. J. Han, J. Pei, and Y. Yin, “Mining frequent patterns without candidategeneration,” in Proc. ACM SIGMOD Int. Conf. Manag. Data,2000, pp. 1–12.
9. J. Han, J. Wang, Y. Lu, and P. Tzvetkov, “Mining top-k frequentclosed patterns without minimum support,” in Proc. IEEE Int.Conf. Data Mining,2002, pp. 211–218.
10. S. Krishnamurthy, “Pruning strategies for mining high utilityitem sets,” Expert Syst. Appl., vol. 42, no. 5, pp. 2371–2381, 2015.
11. [11] C. Lin, T. Hong, G. Lan, J. Wong, and W. Lin, “Efficient updatingof discovered high-utility item sets for transaction deletion indynamic databases,” Adv. Eng. Informat., vol. 29, no. 1, pp. 16–27,2015.
12. G. Lan, T. Hong, V. S. Tseng, and S. Wang, “Applying the maximumutility measure in high utility sequential pattern mining,”Expert Syst.Appl., vol. 41, no. 11, pp. 5071–5081, 2014.
13. Y. Liu, W. Liao, and A. Choudhary, “A fast high utility item setsmining algorithm,” in Proc. Utility-Based Data Mining Workshop,2005, pp. 90–99.
14. M. Liu and J. Qu, “Mining high utility item sets without candidategeneration,” in Proc. ACM Int. Conf. Inf. Knowl. Manag., 2012,pp. 55–64.
15. J. Liu, K. Wang, and B. Fung, “Direct discovery of high utilityitem sets without candidate generation,” in Proc. IEEE Int. Conf.Data Mining, 2012, pp. 984–989.
16. Y. Lin, C. Wu, and V. S. Tseng, “Mining high utility item sets in bigdata,” in Proc. Int. Conf. Pacific-Asia Conf. Knowl. Discovery DataMining, 2015, pp. 649–661.
17. Y. Li, J. Yeh, and C. Chang, “Isolated items discarding strategy fordiscovering high-utility item sets,” Data Knowl. Eng., vol. 64, no. 1,pp. 198–217, 2008.
18. G. Pyun and U. Yun, “Mining top-k frequent patterns with combinationreducing techniques,” Appl. Intell., vol. 41, no. 1, pp. 76–98,2014.
19. T. Quang, S. Oyanagi, and K. Yamazaki, “ExMiner: An efficientalgorithm for mining top-k frequent patterns,” in Proc. Int. Conf.Adv. Data Mining Appl., 2006, pp. 436 – 447.