Temporal Fuzzy Utility Mining with Upper-Bound

Wei-Ming Huang[a], Guo-Cheng Lan[b], Tzung-Pei Hong[a,c] and Ming-Chao Chiang[a]

[a]Department of Computer Science and Engineering,
National Sun Yat-sen University, Kaohsiung 804, Taiwan
[b]Education Solutions and Service (ESS) Department
Delta Electronics, Inc., Taipei, 114, Taiwan
[c]Department of Computer Science and Information Engineering,
National University of Kaohsiung, Kaohsiung 811, Taiwan
granthill168@gmail.com, rrfoheiay@gmail.com, *tphong@nuk.edu.tw,
mcchiang@cse.nsysu.edu.tw

Abstract

Fuzzy utility mining reflects fuzzy degrees of quantities and profits for high utility itemsets. In generally, transaction time is also concerned, and not all products sold are always on the shelf. Thus, in this paper we present an effective framework, which considers the transaction period of each product from the first transaction it appears to the last transaction in the whole database for mining temporal high fuzzy utility itemsets. For reflecting the downward-closure property, an effective upper-bound model is designed to avoid information loss in mining. Based on the proposed model, the efficient two-phase algorithm named TP-THFU is developed to deal with temporal fuzzy utility mining problem.

1 Introduction

Data mining, known as knowledge discovery in database, becomes more and more important technology for extracting particular patterns. Traditional association-rule mining was first proposed to look for the relationship of items from a set of data [1][2]. In retail market, a transaction usually included not only items (or products) but also the quantities, price, or costs of the items. Thus, traditional association-rule mining techniques cannot be capable of dealing with quantities of items. As a result of the reason, Srikant *et al*. then proposed a new issue named quantitative association rule mining [8]. However, it is difficult task to decide what the suitable intervals are in each attribute and the mined rules are not to be understood easily by users.

Chan *et al*. thus proposed another issue, namely utility mining, which considered not only quantities of items but also profits of items in transaction database to find itemsets with high utility values in databases [4]. However, downward-closure property in association-rule mining cannot be kept in the utility mining problem since the utility of an itemset may be larger than, smaller than, or equal to that of its subsets when its itemset length increases. To overcome this problem, Liu *et al*. then presented the two-phase utility mining algorithm, which decomposed two main phases to find high utility itemsets in quantitative transaction databases [7]. However, most approaches [11][12] in fuzzy utility mining are designed for the traditional quantitative databases, and thus they cannot be directly applied for mining high fuzzy utility itemsets in temporal quantitative transaction databases with the consideration of transaction time stamps because of the high complexity computation.

To overcome this problem, temporal fuzzy utility mining algorithm is proposed, to consider not only the quantities and profits of items in a transaction but also the transaction occurring time of items and the minimum operator in fuzzy set theory to find temporal high utility itemsets in databases. With our proposed temporal-based fuzzy utility function, the actual fuzzy utility of an itemset can be evaluated. To avoid information loss in mining, besides, both the temporal-based fuzzy utility upper-bound model and the two-phase temporal fuzzy utility mining approach are developed to achieve this goal.

The rest parts in this paper are organized as follows. The related works are reviewed in Section 2. The problem to be solved and definitions are described in Section 3. The execution details of the proposed approach are explained in Section 4. The experimental result is showed in Section 5. Finally, conclusions are discussed in Section 6.

2 Related Works

Different from traditional association-rule mining [1][2], transactions usually contain some information, like quantity, profit, and cost of an item. However, some high-profit but low-frequency products may not be found by association-rule mining techniques. To solve this problem, Chan *et al*. presented a new issue named utility mining, which thought about both the quantity and profit to calculate the actual utility value of items for high utility itemsets [4]. The utility table of items is given by user's expectation. According to both transaction database and utility table, the discovered itemset is able to better match a user's expectations.

However, the downward-closure property in association-rule mining cannot be kept in the utility mining problem. Liu *et al*. then proposed the transaction-weighted utilization model to avoid information loss in mining. In addition, the two-phase approach was presented to deal with the utility mining problem. Afterward, several studies related to utility mining have also been published [3][5][6][10][11].

Traditional association-rule mining techniques are not capable of handling quantitative values of items [3]. To overcome the problem, Srikant *et al*. proposed a new research issue, namely quantitative association rule mining, which the domain value of each attribute is divided into several range values, to find quantitative association rules [3]. However, there exist some weakness on how to determine the suitable value ranges for the domain values of each attribute, and also it is not easily to be comprehended by users.

Based on the above reasons, Kuok *et al*. first proposed a new research issue, fuzzy data mining [16]. The main concept behind their study is that the quantitative values of items are converted into linguistic regions by the fuzzy set theory, and a minimum operator in fuzzy theory was applied to obtain the overlap value (minimum value) of membership regions in different items. The interesting knowledge with simplicity and comprehensibility for fuzzy data mining could be found from the set of transactions with linguistic regions. Then, Hong *et al*. proposed an effective *Apriori*-based mining algorithm, which adopted a minimum operator in fuzzy theory to count the scalar cardinality value for an itemset in a transaction database, to find interesting fuzzy association rules [14]. In addition, Hong *et al*. proposed an advanced mining approach, which considered the trade-off problem between number of rules and the cost of computation time. The main concept is that only the fuzzy term with the highest fuzzy count for the items could be kept in the set of frequent fuzzy 1-itemsets. Hence, a great deal of candidates could thus be avoided in terms of finding frequent fuzzy itemsets, and also the execution efficiency could be improved [15].

As mentioned previously, however, how to get suitable quantities of products in a product combination is quite important. To address this, Wang *et al*. proposed a new research issue, namely fuzzy utility mining, which integrated fuzzy set theory into utility mining, to find high fuzzy utility itemsets in quantitative databases [9]. In Wang *et al*.'s study [9], the new fuzzy utility function was presented to evaluate the fuzzy utility of an item by the corresponding linguistic region value and degree value in the membership function of that [13]. Lan *et al.* then proposed another new fuzzy utility function, which considers not only quantities and profits of items but also the minimum operator principle of fuzzy set theory, to evaluate the actual fuzzy utility of an itemset in a set of transactions [11]. Like traditional utility mining, in addition, the downward-closure property cannot be kept in fuzzy utility mining, and thus the effective fuzzy-utility upper-bound was designed to avoid any information loss in fuzzy utility mining. In addition, the two-phase fuzzy utility mining approach was presented to cope with the problem of finding high fuzzy utility itemsets. However, most of the studies related to fuzzy utility mining do not consider the exhibition time periods of items from the first transaction time of the items to the last transaction time in the database.

3 Problem Statement and Definitions

To explain the proposed method, assume a temporal quantitative transaction database (*TQD*) is given in Table 1. There are five distinct items in the transactions and the value attached to each item in the corresponding slot is the quantity sold in a transaction. Also, assume the profit value of each item is shown in Table 2, and the same membership functions with the two fuzzy regions, *Low* and *High*, are given for the five items.

Table 1: The set of six quantitative transactions

Period	*TID*	*A*	*B*	*C*	*D*	*E*
P1	$Trans_1$	5	0	4	0	0
P2	$Trans_2$	3	3	0	0	0
P3	$Trans_3$	0	0	1	0	0
P4	$Trans_4$	3	2	4	4	1
P5	$Trans_5$	3	5	0	0	2
P6	$Trans_6$	3	0	0	4	2

Table 2: The profit values of the five items

Item	*Profit*
A	1
B	8
C	5
D	3
E	6

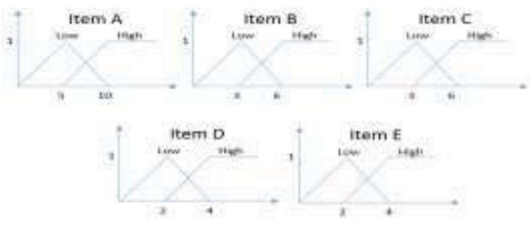

Figure 1. Membership functions for the five items

According to the above example, a set of terms related to the proposed temporal utility fuzzy mining (abbreviated as *TUFM*) is then defined as follows.

Definition 1. Let $T = \{t_1, t_2, ..., t_i, ...\}$ be a set of mutually disjoint time periods, where t_i denotes the i-th time period in the complete periods, T.

Definition 2. An itemset X is a subset of items; That is, $X \subseteq I$. If $|X| = r$, the itemset X is called an r-itemset. For example, the itemset $\{AB\}$ contains 2 items and is called a *2*-itemset.

Definition 3. A quantitative transaction (*Trans*) is composed of a set of purchased items with their quantities. For example in Table 1, the second quantitative transaction contains the two items and their quantities are all 3.

Definition 4. A temporal quantitative transaction database *TQD* is composed of a set of quantitative transactions. That is, $TQD = \{Trans_1, Trans_2, …, Trans_y, …, Trans_z\}$, where $Trans_y$ is the y-th quantitative transaction and z is the number of transactions.

Definition 5. The quantitative value, v_{yz}, is the quantity of the z-th item i_z in a transaction $Trans_y$. For example in Table 1, $v_{2,B} = 1$.

Definition 6. The fuzzy set f_{yz} of the quantitative value v_{yz} of the z-th item i_z in a $Trans_y$ can be represented by the given membership functions for the item i_z as:

$$f_{yz} = \left(\frac{f_{yz1}}{R_{z1}} + \frac{f_{yz2}}{R_{z2}} + ... + \frac{f_{yzl}}{R_{zl}} + ... + \frac{f_{yzh}}{R_{zh}} \right)$$

where h is the number of regions for the item i_z, R_{zl} is the l-th fuzzy region (linguistic term) of i_z, and f_{yzl} is the fuzzy membership value of v_{yz} of i_z in the l-th fuzzy region R_{zl}.

For example in Table 1, the quantitative value of item B in $Trans_2$ can be converted to

$$f_{2,B} = \left(\frac{1.0}{B.Low}, \frac{0}{B.High} \right)$$

by using the given membership functions of item B shown in Figure 1.

Definition 7. The external utility $s(i)$ of an item i is the corresponding utility value of the item in a utility table. For example, the external utility $s(C)$ of the item C is 5 from Table 2.

Definition 8. The fuzzy utility fu_{yzl} of the l-th fuzzy region of an item i_z in a transaction $Trans_y$ is the external utility $s(i_z)$ of the item i_z multiplied by the quantity value v_{yz} and the fuzzy membership value f_{yzl} of v_{yz} in the l-th fuzzy region R_{yzl}. That is,

$$fu_{yzl} = f_{yzl} * v_{yz} * s(i_z)$$

For example from Tables 1 and 2, the quantity and the profit of the item B in $Trans_2$ are 3 and 8, respectively, and the fuzzy membership value of the quantity of item B in the first fuzzy region ($R_{B.Low}$) is 1 according to the given membership functions of item B (as shown in Figure 1). Thus, $fu_{2,B.Low} = 1*3*8 = 24$.

Definition 9. The transaction utility fuzzy tfu_y of a quantitative transaction $Trans_y$ is the summation of the fuzzy utility values of all the items in $Trans_y$. That is,

$$tfu_y = \sum_{i_z \subseteq Trans_y} fu_{yz}$$

where fu_{yz} is the fuzzy utility of the z-th item i_z in $Trans_y$. For example in $Trans_5$, $tfu_5 = fu_{5,\{A.Low\}} + fu_{5,\{B.Low\}} + fu_{5,\{B.High\}} + fu_{5,\{E.Low\}} = 1.8 + 13.2 + 26.4 + 12 = 53.4$.

Definition 10. The fuzzy utility fu_{yX} of a fuzzy itemset X in a transaction $Trans_y$ is the summation of the fuzzy utility values of all the fuzzy terms in X in $Trans_y$. That is,

$$fu_{yX} = f_{yX} * \sum_{R_{zl} \subseteq X} v_{yz} * s(i_z),$$

where v_{yz}, $s(i_z)$, and f_{yzl} represent the quantity of item i_z in $Trans_y$, the profit value of i_z, and the membership value of X, respectively. The last term can be calculated by

$$\underset{R_{zl} \in X}{Min} f_{yzl}$$

, meaning the minimum of all the membership values of all the fuzzy terms in X in $Trans_y$ by using the minimum operator.

For example in Table 1, the membership values of the two fuzzy terms *A.Low* and *C.Low* in the fuzzy itemset $\{A.Low, C.Low\}$ in $Trans_1$ are 1 and 0.66, respectively. By using the minimum operator, the membership value for the *2*-itemset $\{A.Low, C.Low\}$ in $Trans_1$ is thus 0.66. The fuzzy utility of the fuzzy *2*-itemset $\{A.Low, C.Low\}$ in $Trans_1$ can thus be calculated as 0.66*((5*1) + (4*5)), which is 16.5.

Definition 11. Let the start transaction period STP_{iz} of the item i_z be the corresponding time period of the first transaction occurring time of the item i_z in *TQD*. For example in Table 1, $STP_A = P1$ and $STP_D = P4$, respectively.

Definition 12. Let the maximal common transaction period of an itemset X, $MCTP_X$, be the combination of the common transaction periods of all items in an itemset X. For example, if the itemset $\{AD\}$ is composed of item A and item D, the maximal common transaction period $MCTP_{\{AD\}}$ of $\{AD\}$ is represented as $MCTP_{\{AD\}} = MCTP_{\{A\}} \cap MCTP_{\{D\}} = P4$.

Definition 13. The temporal fuzzy utility ratio $tfur_X$ of a fuzzy itemset X in TQD is the summation of all fuzzy utility values of X in all the transactions including X over the summation of the transaction fuzzy utility of all transactions in the maximal common transaction period $MCTP_X$ of X. That is,

$$tfur_X = \sum_{X \in Trans_y \cap Transy \in MCTP_X} fu_{yX} \Big/ \sum_{Transy \in MCTP_X} tfu_y$$

where fu_{yX} is the fuzzy utility of the fuzzy itemset X in $Trans_y$.

For example in Table 1, since the maximal common transaction period $MCTP_{\{B\}}$ of the itemset $\{B\}$ is from the second transaction $Trans_2$ to the last transaction $Trans_6$ in TQD, the summation of the transaction fuzzy utility values of the five transactions in $MCTP_{\{B\}}$ can be calculated as 153.81 (= 25.8 + 1.65 + 47.16 + 53.4 + 25.8). In addition, item B appears in the three transactions ($Trans_2$, $Trans_4$, and $Trans_5$), and then the fuzzy utility of $\{B\}$ can be calculated as 47.76 (=24 + 10.56 + 13.2). The temporal fuzzy utility ratio $tfur_{\{B\}}$ can be calculated as 47.76 / 153.81, which is 31.05%.

Definition 14. Let λ be a pre-defined minimum temporal fuzzy utility threshold. A fuzzy itemset X is called a temporal-based high fuzzy utility itemset (abbreviated as $THFU$) if $tfur_X \geqq \lambda$. For example in Table 1, if $\lambda = 25\%$, then the fuzzy itemset $\{B.Low\}$ is a temporal-based high fuzzy utility *1*-itemset, $THFU$.

Based on the above definitions, a temporal-based fuzzy utility itemset considers not only the individual profits and quantities of items in TQD but also the membership values of quantities and maximal common transaction periods ($MCTP$) of items. However, the fuzzy utility mining problem doesn't have the downward-closure property of traditional association-rule mining. For example in Table 1, assume the temporal fuzzy utility threshold λ is set at 25%. Since $tfur_{\{A.Low\}} = 6.84\% < 25\%$, $\{A.Low\}$ is not a $THFU$. However, its superset $\{A.Low, B.Low\}$ is a $THFU$ because its $tfur(\{A.Low, B.Low\})$ is 27.24%, larger than λ. This example thus shows the downward-closure property doesn't hold for the problem. The proposed temporal-based fuzzy utility mining is thus much harder than the traditional fuzzy utility mining.

To deal with this, we propose an effective upper-bound model, which is named temporal fuzzy utility upper-bound (abbreviated as $TFUUB$), to avoid information loss. A set of related terms in the proposed $TFUUB$ model is stated below.

Definition 15. The maximal fuzzy utility mfu_{yz} of an item i_z in $Trans_y$ is defined as follows: $mfu_{yz} = max\{fu_{yz1}, fu_{yz2}, \ldots, fu_{yzl}\}$, where fu_{yzl} is the fuzzy utility value of the l-th fuzzy region R_{zl} of the item i_z in $Trans_y$. For example, according to Table 1 and Figure 1, $fu_{5,\{B.Low\}} = 0.33*5*8 = 13.2$ and $fu_{5,\{B.High\}} = 0.66*5*8 = 26.4$. Thus, $mfu_{2,\{B.Low\}} = 26.4$.

Definition 16. The maximal transaction fuzzy utility $mtfu_y$ of a transaction $Trans_y$ is the summation of the maximal fuzzy utility values of all the items in $Trans_y$. That is,

$$mtfu_y = \sum_{i_z \subseteq Trans_y} mfu_{yz} ,$$

where mfu_{yz} is the maximal fuzzy utility of the z-th item i_z in $Trans_y$. For example in $Trans_5$, $mtfu_5 = mfu_{5,\{A.Low\}} + mfu_{5,\{B.High\}} + mfu_{5,\{E.Low\}} = 1.8 + 26.4 +12 = 40.2$.

Definition 17. The minimal transaction period MTP of the transaction periods of all items in the set I is defined as follows: $MTP = min\{TP(i_1), TP(i_2), \ldots, TP(i_z)\}$,where $TP(i_1)$ is the transaction period of the item i_z.

For example in Table 1, the transaction periods of both item D and item E are minimal, the minimal transaction period in Table 1 from $Trans_4$ to $Trans_6$.

Definition 18. The temporal-based fuzzy utility upper-bound ratio $tfuubr_X$ of a fuzzy itemset X is the summation of the maximal transaction fuzzy utility values of all the transactions including X in TQD over the summation of all the transactions in MTP. That is,

$$tfuubr_X = \sum_{X \in Trans_y \cap Transy \in MCTP_X} mtfu_{yX} \Big/ \sum_{Transy \in MTP} tfu_y .$$

For example in Table 1, $tfuubr_{\{A.Low\}} = (mtfu_1 + mtfu_2 + mtfu_4 + mtfu_5 + mtfu_6) / (tfu_4 + tfu_5 + tfu_6) = (18.2 + 25.8 + 1.65 + 40.56 + 40.2 + 25.8) / (47.16 + 53.4 + 25.8) = 152.21 / 126.36$, which is 120.4% .

Definition 19. Let λ be a pre-defined minimum temporal fuzzy utility threshold. A fuzzy itemset X is called a temporal-based high fuzzy utility upper-bound itemset (abbreviated as $THFUU$) if $tfuubr_X \geqq \lambda$. For example in Table 1, the fuzzy itemset $\{A.Low\}$ is a temporal-based high fuzzy utility upper-bound *1*-itemset, $THFUU$.

Based on above definitions, the problem to be solved is to find all the temporal-based fuzzy itemsets with their actual fuzzy utility values being larger than or equal to a predefined minimum fuzzy utility threshold under their maximal common transaction periods (MCTPs). To solve the problem, we present a two-phase temporal fuzzy utility mining algorithm (abbreviated as $TP\text{-}TFU$) is also proposed to effectively and efficiently discover temporal-based high fuzzy utility itemsets from *TQD*.

4 The proposed approach

The details of the proposed *TP-THFU* Algorithm (Two-Phase Algorithm for Mining Temporal High Fuzzy Utility Itemsets) are given as follows.

INPUT: A temporal quantitative database *TQD* with *n* quantitative transactions, each of which consists of transaction identification, transaction occurring time, purchased items with quantities, *m* items in *TQD*, each with a membership function, *j* desired time periods, and a predefined minimum temporal fuzzy utility threshold *MinTFutil*.

OUTPUT: A final set of temporal high fuzzy utility itemsets (*THFUs*).

Phase I: Finding *THFUUBs*

STEP 1. Transform the occurring time of each transaction in the temporal quantitative database *TQD* into the corresponding time period.

STEP 2. Find the corresponding start period of each possible item i_z in *TQD* according to the first transaction occurring time of the item i_z, and define the start transaction periods of all the items as STP_{iz}.

STEP 3. Convert the quantitative value v_{jyz} of each item i_{jyz} in each temporal quantitative transaction $Trans_{jy}$ within each time period t_j to a fuzzy set f_{jyz}, and the fuzzy set can be represented as:

$$\left(\frac{f_{jyz1}}{R_{jyz1}}+\frac{f_{jyz2}}{R_{jyz2}}+\ldots+\frac{f_{jyzh}}{R_{jyzh}}\right),$$

using the given membership functions for item quantities, where *h* is the number of regions for i_{jyz}, R_{jyzl} is the *l*-th fuzzy region (linguistic term) of i_{jyz}, and f_{jyzl} is the fuzzy membership value of v_{jyz} in region R_{jyzl}.

STEP 4. Initialize the period fuzzy utility *PFU* table as an empty table, in which each tuple consists of two fields: periodical identification and period fuzzy utility of the period.

STEP 5. For each transaction $Trans_y$ in *TQD*, do the following substeps:

(a) Find the fuzzy utility fu_{yzl} of the *l*-th fuzzy region of item i_z in $Trans_y$ as follows:

$$fu_{yzl} = f_{yzl} * v_{yz} * s(i_z).$$

(b) Find the maximal fuzzy utility mfu_{yz} of each item i_z in $Trans_y$. That is, $mfu_{yz} = max\{fu_{yz1}, fu_{yz2}, \ldots, fu_{yzl}\}$, where fu_{yzl} is the fuzzy utility value of the *l*-th fuzzy region R_{zl} of the item i_z in $Trans_y$.

(c) Calculate the maximal transaction fuzzy utility value $mtfu_y$ of $Trans_y$. That is,

$$mtfu_y = \sum_{i_z \subseteq Trans_y} mfu_{yz},$$

where mfu_{yz} is the maximal fuzzy utility mfu_{yz} of the *z*-th item i_z in $Trans_y$.

(d) Find the transaction fuzzy utility tfu_y of each $Trans_y$ as follows:

$$tfu_y = \sum_{i_z \subseteq Trans_y \cap Trans_y \subseteq TQD} fu_{yzl}.$$

(e) Add the transaction fuzzy utility tfu_y of $Trans_y$ into the corresponding period fuzzy utility field value of the period in the *PFU* table.

STEP 6. Initialize the temporary *1*-itemset (TI_1) table as an empty table, in which each tuple consists of two fields: fuzzy itemset and its temporal fuzzy-utility upper-bound.

STEP 7. For each $Trans_y$ in *TQD*, check the item *i* in $Trans_y$ whether or not it has existed in the *TI* table. If yes, only add the maximal transaction fuzzy utility $mtfu_y$ of $Trans_y$ to its corresponding temporal fuzzy-utility upper-bound field value in the TI_1 table; otherwise, put the item *i* and add its $mtfu_y$ value to the corresponding temporal fuzzy-utility upper-bound field value in the TI_1 table.

STEP 8. Find the minimal transaction period *MTP* of the transaction periods of all the items in *TQD* by their STP_{iz}.

STEP 9. For each item *i* in the TI_1 table, do the following substeps:

(a) Calculate the temporal fuzzy-utility upper-bound ratio $tfuubr_i$ of *i* by the minimal transaction period *MTP* as follows:

$$tfuubr_i = \frac{tfuub_i}{\sum_{j \in MTP} \sum_{Trans_{jy} \in TQD_j} mtfu_{jy}}.$$

(b) Check $tfuubr_i$ of the item *i* whether it satisfies the minimum temporal fuzzy utility threshold (*MinTFUtil*). If it is, the item *i* is put the set of temporal high fuzzy-utility upper-bound 1-itemsets $THFUUB_1$; otherwise, omit it.

STEP 10. Set *r* = 1, where *r* represents the number of items in the current itemset to be processed.

STEP 11. Initialize the temporary itemset $TI_{(r+1)}$ table as an empty table, in which each tuple consists of two fields: fuzzy itemset and its temporal fuzzy-utility upper-bound.

STEP 12. Join r-itemsets in the set of $THFUUB_r$ to generate all possible (r+1)-itemsets by using a similar way used in the *Apriori* algorithm, and put these (r+1)-itemsets in the temporary itemset $TI_{(r+1)}$ table. If no candidate itemsets are generated from the set of $THFUUB_r$, do STEP 17; otherwise, do the next step.

STEP 13. Check each candidate (r+1)-itemest X in the $TI_{(r+1)}$ table whether its each r-sub-itemset exists in the set of $THFUUB_r$. If one of its all r-sub-itemsets does not exist in the set of $THFUUB_r$, the candidate itemset X is removed from the $TI_{(r+1)}$ table; otherwise, it is kept in the $TI_{(r+1)}$ table.

STEP 14. Scan the set of transactions in *TQD* to find the required the temporal fuzzy-utility upper-bound value of each itemset X in the $TI_{(r+1)}$ table by the maximal transaction fuzzy utility values of the transactions including X in *TQD*.

STEP 15. For each itemset X in the $TI_{(r+1)}$ table, do the following substeps:

(a) Calculate the temporal fuzzy-utility upper-bound ratio $tfuubr_X$ of X by the minimal transaction period *MTP* as follows:

$$tfuubr_X = \frac{tfuub_X}{\sum_{j \in MTP} \sum_{Trans_{jy} \in TQD_j} mtfu_{jy}}.$$

(b) Check the $tfuubr_X$ value of the itemset X whether it satisfies the minimum temporal fuzzy utility threshold (*MinTFUtil*). If it is, X is put the set of temporal high fuzzy-utility upper-bound *(r+1)*-itemsets $THFUUB_{(r+1)}$; otherwise, omit it.

STEP 16. If the set $THFUUB_{(r+1)}$ is not empty, then set $r = r + 1$ and repeat STEPs 12 to 17; otherwise, do the next step.

Phase II: Finding *THFUs*

STEP 17. Put the temporal high fuzzy utility itemsets from all the *THFUUB* sets in the final set, *THFUUBs*.

STEP 18. Scan the set of transactions in *TQD* to find the required the fuzzy utility value of each itemset X in the set *THFUUBs*.

STEP 19. For each itemset X in the $TI_{(r+1)}$ table, do the following substeps:

(a) Find the maximal common transaction period $MCTP_X$ of the start transaction periods of *r+1* items in X according to the start transaction period information *STP* of the items as follows: $MCTP_X = max\{STPi_1, STPi_2, \ldots, STPi_{(r+1)}\}$, where $STPi_{(r+1)}$ represents the (r+*1*)-th item $i_{(r+1)}$ in X.

(b) Find the total periodical fuzzy utility $tpfu_X$ of X by its $MCTP_X$ from the *PFU* table as follows:

$$tpfu_X = \sum_{j \in MCTP_X} pfu_j.$$

(c) Calculate the temporal fuzzy-utility ratio $tfur_X$ of X by the maximal common transaction period $MCTP_X$ of X as follows:

$$tfur_X = \frac{fu_X}{tpfu_X}.$$

(d) Check $tfur_X$ of the itemset X whether it satisfies the minimum temporal fuzzy utility threshold (*MinTFUtil*). If it is, X is put the set of temporal high fuzzy-utility upper-bound *(r+1)*-itemsets $THFUUB_{(r+1)}$; otherwise, omit it.

STEP 20. Output the final set of temporal high fuzzy utility itemsets, *THFUs*.

5 Experimental Evaluation

In this section, the proposed *TP-THFU* algorithm was implemented in J2SDK 1.8.0 and executed on a computer with 2.3 GHz CPU and 4 GB memory. Experiments were made on the synthetic T10I4N4KD200K dataset to evaluate the performance of the proposed approach. Note that the execution time of the proposed *TP-THFU* is for finding *THFU1* and *THFU2* in the experiments. The experiments are described to show execution times on the dataset with the minimum temporal fuzzy utility threshold varying from 1.00% to 5.00%. The results are shown in Figure 2.

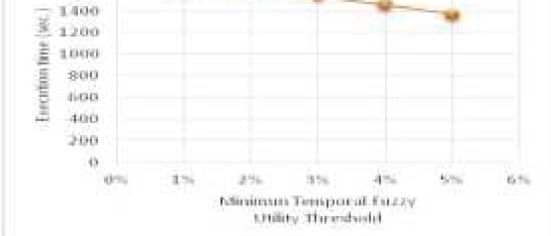

Figure 2. Performance of *TP-THFUs*

6 Conclusions

We have presented a new temporal fuzzy utility function and an efficient Two-Phase Algorithm, named *TP-THFU*, to find temporal high fuzzy utility itemsets in temporal quantitative transaction databases. In addition, we propose an effective fuzzy utility upper-bound model to keep the downward-closure property in temporal fuzzy utility mining problem. Experimental results also show the proposed approach has good performance in terms of execution efficiency.

References

[1] R. Agrawal and R. Srikant, "Fast algorithm for mining association rules," In *Proceedings of the International Conference on Very Large Data Bases*, pp. 487-499, 1994.

[2] R. Agrawal, T. Imielinksi, and A. Swami, "Mining association rules between sets of items in large database," In *Proceedings of the ACM SIGMOD International Conference on Management of Data,* pp. 207-216, 1993.

[3] Srikant, R., and Agrawal, R.,"Mining Quantitative Association Rules in Large Relational Tables". In *Proceedings of ACM SIGMOD Conference on Management of Data*, pp. 1-12, 1996 ACM Press.

[4] R. Chan, Q. Yang, and Y. D. Shen, "Mining high utility itemsets," In *Proceedings of the 3rd IEEE International Conference on Data Mining*, pp 19-26, 2003.

[5] Raymond Chan, Qiang Yang and Yi-Dong Shen, "Mining high utility itemsets", *Third IEEE International Conference on Data Mining*, pp 19-26. 2003

[6] Ying Liu, Wei-Keng Liao, and Alok Choudhary, "A fast high utility itemsets mining algorithm," *The Utility-Based Data Mining Workshop*, pp. 90-99, 2005.

[7] Y. Liu, W. K. Liao, and A. Choudhary, "A fast high utility itemsets mining algorithm," In *Proceedings of the Utility-Based Data Mining Workshop*, pp. 90-99, 2005.

[8] R. Srikant and R. Agrawal, "Mining quantitative association rules in large relational tables," In *Proceedings of the ACM SIGMOD Conference on Management of Data*, pp. 1-12, 1996.

[9] C. M. Wang, S. H. Chen, and Y. F. Huang, "A fuzzy approach for mining high utility quantitative itemsets," In *Proceedings of the IEEE International Conference on Fuzzy Systems*, pp. 20-24, 2009.

[10] Yeong-Chyi Lee, Tzung-Pei Hong and Tien-Chin Wang, "Multi-level fuzzy mining with multiple minimum supports", *Expert Systems with Applications*, Vol. 34, pp. 459-468, 2008- Elsevier.

[11] G. C. Lan, T. P. Hong, Y. H. Lin, and S. L. Wang, "Fuzzy Utility Mining with Upper-Bound Measure", *Applied Soft Computing*, Vol. 30, pp. 767-777, 2015.

[12] Chia-Ming Wang, Shyh-Huei Chen, and Yin-Fu Huang "A Fuzzy Approach for Mining High Utility Quantitative Itemsets", *Fuzzy-IEEE 2009,* pp. 20-24, 2009.

[13] Guo-Cheng Lan, Tzung-Pei Hong, Yi-Hsin Lin, and Shyue-Liang Wang "Fast discovery of high fuzzy utility itemsets", Systems, Man and Cybernetics, pp. 2764-2767, 2014.

[14] T. P. Hong, C. S. Kuo, and S. C. Chi, "A fuzzy data mining algorithm for quantitative values", In *Proceedings of the 3rd International Conference on Knowledge-Based Intelligent Information Engineering System*, pp. 480-483, 1999.

[15] T. P. Hong, C. S. Kuo, and S. C. Chi, "Trade-off between computation time and number of rules for fuzzy mining from quantitative data", *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 9, No. 5, pp. 587-604, 2001.

[16] C. M. Kouk, A. Fu, and M. H. Wong, "Mining fuzzy association rules in database," *ACM SIGMOD Record*, Vol. 27, No. 1, pp. 41-46, 1998.