

DISCOVERING IMPERCEPTIBLE ASSOCIATIONS BASED ON INTERESTINGNESS: A UTILITY-ORIENTED DATA MINING APPROACH

S Shankar^{1*} and *T Purusothaman*²

¹*Department of Information Technology, Sri Krishna College of Engineering and Technology, Coimbatore, Tamilnadu, India.*

Email: shanx80@yahoo.in

²*Department of Computer Science and Engineering, Government College of Technology, Coimbatore, Tamilnadu, India.*

Email: purushgct@yahoo.com

ABSTRACT

This article proposes an innovative utility sentient approach for the mining of interesting association patterns from transaction databases. First, frequent patterns are discovered from the transaction database using the FP-Growth algorithm. From the frequent patterns mined, this approach extracts novel interesting association patterns with emphasis on significance, utility, and the subjective interests of the users. The experimental results portray the efficiency of this approach in mining utility-oriented and interesting association rules. A comparative analysis is also presented to illustrate our approach's effectiveness.

Keywords: Data Mining, Frequent Patterns, Association Rules, FP-Growth, Economic Utility, Weight, Significance, Interestingness, Subjective Interestingness

1 INTRODUCTION

A noteworthy field of computer science that has attracted great interest for research and development from a huge variety of people is data mining. The principal motivation behind the emergence of data mining stems from decision support problems that have affected a large number of business organizations (Tsur, 1990; Wang et al., 1994). Data mining, also termed Knowledge Discovery in Databases (KDD), has been defined as "The nontrivial extraction of hidden, novel, and potentially useful information from data" (Frawley et al., 1992). Data mining makes use of machine learning and various statistical and visualization techniques so as to determine and represent knowledge in an easily interpretable form (Soundararajan et al., 2005). Data mining tasks can be generally classified into *descriptive mining* and *predictive mining*. The mined information is articulated as a model of the semantic structure of the dataset, where the prediction or classification of the collected data is facilitated by means of the model (Cunningham & Holmes, 1999). Recently, the incorporation of utility constraints into data mining tasks has emerged as an area of intensive research in data mining.

Utility-based data mining (Weiss et al., 2006; Yeh et al., 2007) is an extensive area that focuses on all aspects of economic utility in data mining and is aimed at incorporating utility into both predictive and descriptive data mining tasks. There have been a record number of transactional databases available, with computers and e-commerce gaining fast and widespread recognition. The chief focus of data mining on transactional databases is the mining of association rules that determine the correlation among items in transaction records. Data mining researchers are chiefly concerned with qualitative aspects of attributes (e.g., significance and utility) rather than quantitative ones (e.g., number of appearances in a database, etc), as qualitative properties are essential for exploiting completely the attributes present in the dataset. Within the data mining community, the discovery of interesting association rules has long been recognized as a way of improving the business utility of an enterprise. Business improvement demands the extraction of interesting association patterns that are both statistically and semantically significant to the business utility.

Literature presents a plentiful array of association rule mining algorithms (Kotsiantis & Kanellopoulos, 2006; Aggrawal & Yu, 1998; Coenen & Leng, 2001; Han et al., 2000; Lee et al., 2003; Dong & Tjortjis, 2003; Wang & Tjortjis, 2004; Verma et al., 2005; Yuan & Huang, 2005), a large number of which speed-up association rule mining rather than generate effective rules (Bodon, 2003; Lee et al., 2003; Wang & Tjortjis, 2004). Most techniques for extracting valid and genuine association patterns are quantitative in nature (Srikant & Agrawal,

1996; Kotsiantis & Kanellopoulos, 2006; Dong & Tjortjis, 2003; Aggrawal & Yu, 1998). Qualitative attributes are also essential to completely exploit object attributes present in the data set. Moreover, a certain number of methods are available in the literature for mining weighted association rules (Cai et al., 1998; Wang et al., 2000). In recent times, incorporating utility constraints in itemset and association rule mining has gained enormous popularity (Yeh et al., 2007; Podpecan et al., 2007). We have presented a survey and comparative study of the significant research based on utility for itemset and association rule mining in Shankar and Purusothaman (2009). On the basis of the survey conducted, we have presented a novel utility sentient approach for mining interesting association rules in our earlier work (Shankar & Purusothaman, 2009).

In this article, we present our previous work, novel utility-based approach for the mining of high utility and interesting association patterns from transactional databases, with detailed results based on experimentation and a comparative analysis with Muyebe et al. (2008) and Khan et al. (2008) to emphasize the performance of the our research. The primary objective of our approach is to identify novel and interesting association rules from the historical buying patterns of customers, so as to increase the business of an enterprise. This approach makes use of the FP-Growth algorithm for mining frequent patterns. Subsequently, the mined frequent patterns are subjected to the computation of interestingness measure and utility weight (significance of an item). The process results in a class of novel, utility-oriented, and interesting association patterns. We compare the performance of our approach to two existing weighted association rule mining approaches Weighted ARM (Muyebe et al., 2008) and Weighted Utility ARM (Khan et al., 2008) to illustrate the efficacy of our approach. An experimental and comparative analysis of this approach shows that it identifies interesting patterns that are both statistically and semantically important in improving business utility.

The remainder of the paper is organized as follows: section 2 provides a brief review of the research related to our research. The approach proposed for mining utility sentient and novel interesting association patterns is presented in section 3. Section 4 presents the experimental results of our approach with comparative analysis. Section 5 sums up the paper.

2 REVIEW OF RELATED RESEARCH

A handful of research is available in the literature for mining frequency patterns and association rules based on weight and utility. A brief review of the significant research is presented here.

Aggrawal et al. (1993) have proposed a proficient algorithm that generates all important association rules from the database. Buffer management, novel estimation, and pruning techniques are the options added to the proposed algorithm. The problem of mining association rules from huge relational tables containing both quantitative and categorical attributes has been dealt with by Srikant et al. (1996). Moreover, they have established a measure of partial completeness, which quantifies the information lost due to partitioning.

Aggarwal et al. (1998) have conducted an extensive survey of the research available for association rule mining. They have also discussed numerous variations of the association rule problem proposed in the literature and their practical applications. A method of mining the weighted association rule has been proposed by Cai et al. (1998). They have offered two distinct definitions for weighted support, with normalization and without normalization, and have also proposed a novel algorithm on the basis of the support bounds.

A two-fold approach for association rule mining has been presented by Wang et al. (2000). First, frequent itemsets were generated, and then the maximum weighted association rules were derived using an "ordered" shrinkage approach. Coenen and Leng (2001) have given a method for generating frequent itemsets that minimizes the task by using efficient restructuring of data accompanied by a partial computation of the totals required. Dong and Tjortjis (2003) have discussed an enhanced memory efficient data structure of a quantitative approach to mine association rules from data. They have combined the best features of the three algorithms (the Quantitative Approach, DHP, and Apriori) in their approach. Bodon (2003) has presented an implementation procedure for the Apriori algorithm that exceeded all other available implementations.

An efficient algorithm for association rule mining was presented by Wang and Tjortjis (2004). The enormous time requirement incurred for large itemset generation was substantially reduced by performing a single scan of the database and then employing logical operations to achieve large itemset generation. Verma et al. (2005) have presented an innovative algorithm for discovering association rules on time dependent data utilizing

efficient T-tree and P-tree data structures. The algorithm achieves considerable benefit in terms of time and memory even after integrating the time dimension.

Yuan and Huang (2005) have proposed the Matrix Algorithm for proficient generation of large frequent candidate sets. First, the algorithm generates a matrix containing values 1 or 0 by conducting a single pass over the concrete database. The resulting matrix is employed in the generation of frequent candidate sets. Ultimately, from the frequent candidate sets, association rules are mined. Palshikar et al. (2007) presented the concept of heavy itemsets that compactly symbolizes an exponential number of rules. They have given a proficient theoretical characterization of a heavy itemset and a proficient greedy algorithm for generating an anthology of disjoint heavy itemsets in a given transactional database.

An innovative utility-frequent mining model for the discovery of all itemsets that can generate a user specified utility in transactions has been presented by Yeh et al. (2007). They have proposed two sets of algorithms for efficiently mining utility-frequent item sets: 1) a bottom-up two-phase algorithm (BU-UFM) and 2) a top-down two-phase algorithm (TD-UFM). Podpecan et al. (2007) have presented an innovative and effective algorithm FUFM (Fast Utility-Frequent Mining) that discovers all utility-frequent itemsets within the given utility and supports a constraints threshold.

3 SUGGESTED APPROACH FOR MINING INTERESTING ASSOCIATION PATTERNS

Our approach for interesting association pattern mining is presented in this section. The primary aim of this research is to devise an efficient approach that can mine novel interesting association patterns from a transaction database of an enterprise so as to augment business development. Generally, a transaction database presents a good depiction of the buying patterns of a business's customers. Our approach intends to discover novel interesting association patterns that may or may not be frequent in the database but are likely to improve the enterprise's business. This approach incorporates three factors in addition to frequency for association pattern mining: 1) significance weight, 2) utility, and 3) subjective interestingness to the users. The aforesaid factors were selected for: *item significance* - each item in a transaction has a distinct level of significance based on its importance and *item utility* - each item has a subjective utility, such as profit in dollars or some other utility factor that influences business development.

Interestingness is the measure of the quantity of 'interest' that a pattern evokes upon inspection. Interestingness is considered to be a significant factor in data mining because it influences the extraction of novel and valuable (interesting) patterns and also plays a vital role in determining the efficacy of an enterprise. In general, the measures of interestingness can be categorized into objective and subjective measures. The objective measures of interestingness are commonly represented via statistical or mathematical criteria, while subjective measures deal with more realistic criteria such as efficiency or applicability.

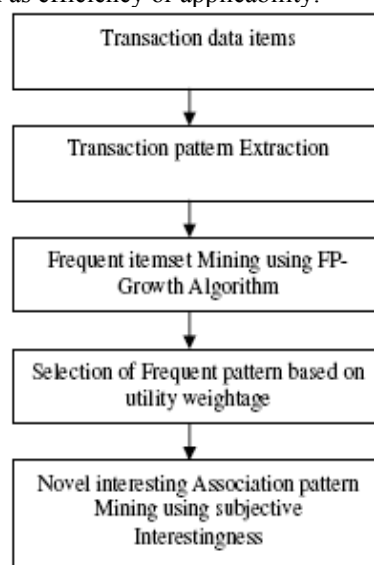


Figure 1. Block diagram of our approach

The above figure depicts the block diagram of the proposed approach. The sequential steps involved in the proposed approach are as follows:

1. Computation of the significance weight for all items based on profit.
2. Determination of frequent itemsets using the FP-Growth algorithm.
3. Determination of the frequent patterns with utility weight greater than a threshold value.
4. Discovery of novel interesting association patterns from the selected patterns on the basis of the interestingness measure.

The input to the proposed approach is a transaction database that contains buying patterns of customers from basket data. The definitions involved in the process of association rule mining from a transaction database D are:

Let $IS = \{i_1, i_2, i_3, \dots, i_m\}$ be a set of items and $TDI = \{t_1, t_2, t_3, \dots, t_n\}$ be a set of transaction data items, where $t_i = \{IS_{i1}, IS_{i2}, IS_{i3}, \dots, IS_{ip}\}$, with $p \leq m$, and $IS_{ij} \in IS$. If $X \subseteq I$ with $k = |X|$ called a k -itemset or simply an itemset. An association rule is of the structure $X \Rightarrow Y$, where X, Y are item sets and $X \cap Y = \Phi$. Two customary measures that serve as the basis for association rule mining are support and confidence.

Support of an itemset is defined as the measure of the number of transactions T containing the itemsets X and Y with respect to TDI .

$$Support(X) = |\{T \in TDI | X \subseteq T\}| / |TDI| \quad (1)$$

Confidence of an association rule $X \Rightarrow Y$ is defined as the ratio of the number of transactions that contain $X \cup Y$ to the number of transactions that contain X .

$$Conf(X \rightarrow Y) = Support(X \cup Y) / Support(X) \quad (2)$$

3.1 Significance weight calculation

The first step of the proposed approach is the computation of the significance weight for all items based on profit. The significance weight corresponding to the items is calculated by

$$SW = P / \sum P_i, i = 1 \text{ to } n, \quad (3)$$

where n = number of items and P = profit of an item.

Subsequently, the data items in the transaction database are represented via a matrix for pattern extraction,

$$T_{ij} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & \dots & t_{1m} \\ t_{21} & t_{22} & t_{23} & \dots & t_{2m} \\ \vdots & & & & \\ t_{n1} & t_{n2} & t_{n3} & \dots & t_{nm} \end{bmatrix}.$$

The patterns found in the transactions are then extracted as a set by making use of the above matrix:

$$T_i = \{\Delta_I(t_{ij}) : t_{ij} \neq 0\}, i = 1, \dots, n \text{ and } j = 1, \dots, m; T_n \ll T_i \quad (4)$$

Where Δ_I = item index of all elements in T_{ij} . The FP-Growth algorithm is then employed to discover all the frequent patterns from the set of extracted transaction patterns T_n .

3.2 FP- Growth algorithm

The FP-growth algorithm is one contemporary approach for mining frequent itemsets (Han et al., 2000). The FP-growth algorithm makes use of a prefix tree representation of the given transaction database (called an FP-

tree). The FP-tree is used to lessen the amount of memory needed for storing transactions. The major steps involved in the FP-growth algorithm are:

- 1) Construction of a memory structure called the FP-tree.
- 2) Repeated application of the actual FP-growth procedure to the constructed FP-tree.
- 3) Discovery of all frequent itemsets by analyzing projections (conditional FP-trees) of the tree in a depth-first manner pertaining to the frequent prefixes mined thus far.

3.3 Frequent pattern selection based on utility weight

The next step in frequent pattern mining is the selection of utility weighted frequent patterns. The set of frequent patterns mined is subjected to utility weight computation. The utility weight of a frequent pattern is computed using equation (5):

$$Wu = \sum_{i=1}^n f * SW_i, \quad (5)$$

where Wu = utility weight,

f = frequency of pattern, and

SW_i = significance weight of the i^{th} item in the current pattern.

Then, based on a predefined threshold α , a set of patterns with higher utility weight is selected:

$$S_p = \{F_p : P(F_p)\} \quad (6)$$

$$P(F_p) = Wu > \alpha,$$

where S_p = set of selected patterns

α = predefined threshold value, and

Wu = utility weight.

3.4 Interesting association pattern mining

Following the utility weight computation, the next step is the mining of novel interesting association patterns from the selected frequent patterns. This sub-section describes the algorithm employed for the mining of these patterns. The subjective interestingness to users is used in the algorithm for the mining of interesting association patterns. The algorithm takes as input: 1) the frequent patterns selected from utility weight and 2) all the frequent patterns mined by the FP-Growth algorithm. The algorithm aims at mining novel interesting association patterns that are present in transactions but may or may not be frequent and likely to increase business. The steps involved in the proposed algorithm for mining interesting association patterns are:

- 1) Extracting all other frequent patterns containing each discrete item in the selected frequent pattern.
- 2) Obtaining the set difference of the selected itemset (pattern) from the extracted itemset.
- 3) Pairing every discrete item in the resulting patterns (set difference) with its reference item.

Thus, for every selected pattern, we obtain a set of patterns, each with two items. The pseudo code for the above operation is given below.

Assumptions:

Sp_{ij} = Individual item in a selected pattern

Sp_i = Individual pattern in a selected pattern set

F_{pg} = Frequent itemsets from the FP-Growth algorithm

RP = Resulting patterns

Pseudo Code:

```

for each selected pattern
  for each item in selected pattern
    for each frequent pattern
      If  $Sp_{ij} \subset Fp_{gi}$ 
         $fp \ll Fp_{gi} \setminus Sp_i$ ;
      end if
       $TP1 \ll S_{p_{ij}} \parallel fp$ ;
    end for
     $TP2 \ll TP1$ ;
  end for
   $RP \ll TP2$  ;
end for

```

- 4) Grouping all patterns from the set of patterns created for a selected pattern with an identical second item, the first item being the reference.
- 5) Computing a single consolidated weight for all the pattern sets obtained in step 4 by adding the individual weights of patterns present. The weight of a pattern is the sum of all of its items' weights. Item weight is the product of the frequency and significance weight of the item. Frequency of an item is defined as the occurrence count of the item in transactions. The frequency of item X in pattern $X \rightarrow Y$ is calculated using equation (7):

$$C = \sum_{i=1}^n tc(X) ; tc(XY) \neq 0, \quad (7)$$

where $tc(X)$ = the count of item X in the current transaction and n = the total number of transactions.

- 6) Computing the interestingness weight of a pattern group by

$$Iw = \sum_{j=1}^m \left(\sum_{i=0}^1 C_i W_i \right), \quad (8)$$

where I_w = interestingness weight and m = number of patterns in the group.

- 7) Combining the second item from each pattern group with its corresponding selected pattern and assigning the interestingness weight of that group to the newly formed pattern.
- 8) Sorting the newly formed patterns on the basis of their interestingness weight. On occurrence of more than one pattern with an identical last item, only the pattern with the highest interestingness weight is used in the subsequent process. The patterns with higher interestingness weight are selected based on a threshold value ψ .

4 EXPERIMENTAL RESULTS AND ANALYSIS

The results and analysis obtained from experimentation on the proposed approach are presented in this section. The proposed approach is programmed in Java. The profit and computed significance weight for individual items are depicted in Table 1. A sample transaction database is shown in Table 2. Tables 1 and 2 illustrate the results of the algorithms described in sections 3.2, 3.3, and 3.4. The frequent patterns discovered by the FP-growth algorithm with their corresponding frequencies are tabulated in Table 3. Table 4 portrays the utility weight of the generated frequent patterns. Table 5 contains the set of patterns with utility weight greater than 2.0. The novel patterns discovered with their interestingness weight are tabulated in Table 6. Table 8 shows the final set of novel interesting patterns selected from Table 7 with a minimum threshold of 10.0.

Table 1. Profit and significance weight of items

Id	Item	Profit	Significance weight
1	A	60	0.2857
2	B	10	4.7619
3	C	30	0.1429
4	D	90	0.4286
5	E	20	9.5239

Table 2. Customer transactions database

TID	A	B	C	D	E
1	10	1	4	1	0
2	0	1	0	3	0
3	2	0	0	1	0
4	0	0	1	0	0
5	1	2	0	1	3
6	1	1	1	1	1
7	0	2	3	0	1
8	0	0	0	1	2
9	7	0	1	1	0
10	0	1	1	1	1

Table 3. Output of the FP-Growth algorithm

Patterns	Frequency
A D	5
D	8
A C D	3
B D	5
A B D	3
C D	4
E D	4
C B D	3
E B D	3
A	5
A C	3
A B	3
C	6
E C B	3
C B	4
E C	3
E B	4
E	5
B	6

Table 4. Utility weight of frequent patterns

Patterns	Frequency	Utility weight
A D	5	3.5714
D	8	3.4286
A C D	3	2.5714
B D	5	2.3809
A B D	3	2.2857
C D	4	2.2857
E D	4	2.0952
C B D	3	1.8571

E B D	3	1.7142
A	5	1.4285
A C	3	1.2857
A B	3	1.0
C	6	0.8571
E C B	3	0.8571
C B	4	0.7619
E C	3	0.7142
E B	4	0.5714
E	5	0.4761
B	6	0.2857

Table 5. Frequent patterns having utility weight > 2.0

Patterns	Frequency	Utility weight
A D	5	3.5714
D	8	3.4286
A C D	3	2.5714
B D	5	2.3809
A B D	3	2.2857
C D	4	2.2857
E D	4	2.0952

Table 6. Interestingness weight of newly formed patterns

Patterns	Interestingness weight
C D A	14.1429
B D A	11.7619
A B D C	10.2380
A D C	8.7143
A C D B	8.4286
E D A	8.1429
D A	8.1428
A D B	6.9047
C D B	4.8095
B D C	4.2380
E D B	4.1428
E D C	3.7142
C D E	3.3809
A C D E	3.3809
D B	3.2857
A B D E	3.2381
B D E	3.2381
D C	2.7143
D E	2.3809
A D E	2.3809

Table 7. Patterns selected from newly formed patterns

Patterns	Interestingness weight
C D A	14.1429
A B D C	10.2381
A C D B	8.4286
C D E	3.3809

Table 8. Novel interesting patterns

Patterns	Interestingness weight
C D A	14.1429
A B D C	10.2381

A discussion of the results obtained from experimentation of our approach is presented here. In general, standard association rule mining algorithms result in enormous patterns, and users are expected to shortlist or select the patterns that are interesting to their own businesses. However, from the results, it is clear that our approach, in contradiction to traditional association rule mining algorithms, generates only a meager number of interesting association patterns that are both statistically and semantically important for business development. Most of the algorithms for weighted and utility association rule mining existing in the literature employ weight and utility measures with support thresholds in frequent pattern mining. These algorithms are likely to miss the itemsets that are of less utility but are of high frequency. This in turn affects business utility because, in most cases, frequency plays a vital part in business development including sales backup and more.

On the whole, different businesses necessitate different levels of interestingness and utility. Therefore, an approach that outputs patterns specific to each criterion is highly valuable to users from different types of businesses. In order to facilitate this specificity, we have incorporated multiple levels of refinement (frequency, significance weight, utility, and interestingness weight), each useful for different types of business. Finally, the ultimate step of our approach, interestingness, is likely to generate novel association patterns by correlating the utility weight and frequent patterns. These novel patterns may or may not be frequent but could be very interesting to users for business development.

4.1 Comparative Analysis

The analysis of the results obtained from the comparison of the proposed approach with standard ARM (Agrawal et al., 1993), weighted ARM (Muyeba et al., 2008), and weighted utility ARM (Khan et al., 2008) is presented in this sub-section. This sub-section provides the comparative results of our approach and a brief description validating the significance of this approach when compared to the existing works of Muyeba et al. (2008) and Khan et al. (2008). Table 9 depicts the different patterns resulting from standard ARM (Agrawal et al., 1993), weighted ARM (Muyeba et al., 2008), and weighted utility ARM (Khan et al., 2008), each with 30% weight, and the proposed approach (high utility and interesting patterns).

Table 9. Comparison Results

#	Standard ARM (Agrawal et al., 1993)	Weighted ARM (Muyeba et al., 2008)	Weighted Utility ARM (Khan et al., 2008)	Proposed Approach	
				High utility patterns (utility weight)	Interesting patterns (interestingness weight)
1	A (50%)	A (30%)	A (0.59)	A → D (3.5)	A → B → D → C (10.2)
2	A → B (30%)	A → B → C → D (38%)	A → C (0.38)	A → B → D (2.2)	A → C → D → B (8.4)
3	A → B → D (30%)	A → B → D (48%)	A → C → D (0.38)	A → C → D (2.5)	C → D → A (14.1)
4	A → C (30%)	A → B → D → E (36%)	A → D (0.590)	B → D (2.3)	C → D → E (3.3)
5	A → C → D (30%)	A → C → D (54%)	B (0.51)	C → D (2.2)	
6	A → D (50%)	A → D (75%)	B → D (0.45)	D (3.4)	
7	B (60%)	A → D → E (34%)	C (0.52)	E → D (2.09)	
8	B → C (40%)	B → C → D (39%)	C → D (0.43)		
9	B → C → D (30%)	B → C → D → E (30%)	D (0.50)		
10	B → C → E (30%)	B → D (50%)	E (0.32)		
11	B → D (50%)	B → D → E (36%)			
12	B → D → E (30%)	C → D (48%)			
13	B → E (40%)	D (72%)			

14	C (60%)	D→E (44%)
15	C→D (40%)	
16	C→E (30%)	
17	D (80%)	
18	D→E (40%)	
19	E (50%)	

From the above table, it is clear that there is a steady decrease in the number of entries in each column of the table. The reduced number of patterns shows that the incorporation of distinct measures (weight, utility, etc.) into association rule mining brings about a progressive refinement of the quality of the discovered patterns. Khan et al. (2008) have shown that weighted utility ARM (WUARM) is better than weighted ARM because WUARM considers all possible patterns and then uses the items' weight and utilities for refinement. The patterns with a highlighted background in column 6 are the novel interesting patterns discovered using the efficient approach presented. An investigation of the patterns discovered by standard ARM, weighted ARM, WUARM, and our approach shows that our approach neglects some significant but invaluable patterns and brings about optimality between weight and utility measures.

For instance, pattern B (0.51) qualified in WUARM is not opted in our approach because even though the pattern is significant, it does not signify to be valuable i.e.) profitable. Also, the high utility patterns of our approach can be seen as an optimal combination of the patterns discovered in weighted ARM and WUARM. Furthermore, the high utility patterns discovered in historical buying patterns certainly signify the importance of the items in the growth of the enterprise. In addition to utility, we have incorporated the subjective interestingness of the users for pattern extraction. The set of novel interesting patterns mined based on the user's subjective interestingness depicts the buying patterns of the future and can serve as an effective means to increase business development.

5 CONCLUSION

Association rules have been utilized extensively to determine customer buying patterns from market basket data. In recent times, researchers have been greatly interested in incorporating utility considerations into association rule mining. Recently, the data mining community has turned to the mining of interesting association rules to facilitate business development by increasing the utility of an enterprise. The above scenario emphasizes the need to discover interesting and utility sentient association patterns that are both statistically and semantically important to business development. In this paper, we have presented a novel approach for mining high utility and interesting association patterns from transaction databases. This approach is aimed at mining association patterns that facilitate improvement of business utility. It focuses on the utility, significance, and interestingness of individual items for the mining of novel association patterns. The mined interesting association patterns are used to offer valuable suggestions to an enterprise for intensifying its business utility. The experimental results and analysis demonstrate the effectiveness of our approach in mining utility sentient and interesting association patterns.

6 REFERENCES

- Aggrawal C., & Yu, P. (1998) Mining large itemsets for association rules. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering* 21(1).
- Agrawal, R., Imielinski, R., & Swami, A. (1993) Mining Associations between Sets of Items in Large Databases. *In proceedings of the ACM SIGMOD Int'l Conf. on Management of data*. Washington, DC.
- Bodon, F. (2003) A Fast Apriori implementation. *In ICDM Workshop on Frequent Itemset Mining Implementations*. Melbourne, Florida, USA.
- Cai, C.H., Fu, A.W-C., Cheng, C. H., & Kwong, W.W. (1998) Mining Association Rules with Weighted Items. *In: Proceedings of 1998 Intl. Database Engineering and Applications Symposium (IDEAS'98)*. Cardiff, Wales, UK.

- Coenen, F., & Leng, P. (2001) Optimizing Association Rule Algorithms Using Itemset Ordering. *Research and Development in Intelligent Systems XVIII: Proc ES2001 Conference*.
- Cunningham, S. J. & Holmes, G., (1999) Developing innovative applications in agriculture using data mining. In the *Proceedings of the Southeast Asia Regional Computer Confederation Conference*. Singapore.
- Dong L., & Tjortjis, C. (2003) Experiences of Using a Quantitative Approach for Mining Association Rules. *Lecture Notes in Computer Science* 2690.
- Frawley, W., Piatetsky-Shapiro, G., & Matheus, C. (1992) Knowledge Discovery in Databases: An Overview. *AI Magazine*.
- Han, J., Pei, J., & Yin, Y. (2000) Mining frequent patterns without candidate generation. In *Proceedings of 2000 ACM SIGMOD International Conference on Management of Data (SIGMOD'00)*. Dallas, TX.
- Khan, S., Mueyba, M. & Coenen, F. (2008) A Weighted Utility Framework for Mining Association Rules. In *proceedings of the Second UKSIM European Symposium on Computer Modeling and Simulation*.
- Kotsiantis, S. & Kanellopoulos, D. (2006) Association Rules Mining: A Recent Overview. *GESTS International Transactions on Computer Science and Engineering* 32 (1).
- Lee, C.-H., Chen, M.-S., & Lin, C.-R. (2003) Progressive Partition Miner: An Efficient Algorithm for Mining General Temporal Association Rules. *IEEE Transactions on Knowledge and Data Engineering* 15(4).
- Mueyba, M., Khan, S., & Coenen, F. (2008) Fuzzy Weighted Association Rule Mining with Weighted Support and Confidence Framework. *New Frontiers in Applied Data Mining: PAKDD 2008 International Workshops (ALSIP)*. Osaka, Japan.
- Palshikar, GK., Kale, MS., Apte, MM. (2007) Association rules mining using heavy itemsets. *Data and Knowledge Engineering* 61(1).
- Podpecan, V, Lavrac, N. & Kononenk, I. (2007) A fast algorithm for mining utility-frequent itemsets. *The 11th European Conference on Principles and Practice of Knowledge Discovery in Databases*.
- Shankar, S. & Purusothaman, T. (2009) A Novel Utility Sentient Approach for Mining Interesting Association Rules. *International Journal of Engineering and Technology (IJET)* 1(5).
- Shankar, S. & Purusothaman, T. (2009) Utility Sentient Frequent Itemset Mining and Association Rule Mining: A Literature Survey and Comparative Study. *International Journal of Soft Computing Applications (IJSCA)* (4).
- Soundararajan, E., Joseph, J.V.M., Jayakumar, C., & Somasekharan, M. (2005) Knowledge Discovery Tools and Techniques. In *Proceedings of the Conference on Recent Advances in Information Technology (READIT)*. Kalpakkam, Tamil Nadu.
- Srikant, R. & Agrawal, R., (1996) Mining Quantitative Association Rules in Large Relational Tables. In *Proc. of ACM SIGMOD Conf. on Management of Data*.
- Tsur, S. (1990) Data dredging. *IEEE Data Engineering Bulletin* 13(4).
- Verma, K., Vyas, O.P., & Vyas, R. (2005) Temporal Approach to Association Rule Mining Using T-Tree and P-Tree. *Lecture Notes in Computer Science* 3587.
- Wang, J. T-L., Chirn, G-W., Marr, T. G., Shapiro, B., Shasha, D., & Zhang, L. (1994) Combinatorial pattern discovery for scientific data: some preliminary results. In *Proceedings of the 1994 ACM SIGMOD International Conference on Management of Data* Minneapolis, MN.
- Wang, W., Yang, J., & Yu, P. S., (2000) Efficient Mining of Weighted Association Rules (WAR). In *Proceedings of the KDD*. Boston, MA.

Wang, C. & Tjortjis, C. (2004) PRICES: An Efficient Algorithm for Mining Association Rules. *Lecture Notes in Computer Science* 3177.

Weiss G., Zadrozny B., & Saar-Tsechansky, M. (2006) Utility-based data mining. *2006 workshop report. SIGKDD Explorations* 8(2).

Yeh, J.-S., Li, Y.-C., & Chang, C.-C. (2007) Two-Phase Algorithms for a Novel Utility-Frequent Mining Model. *Lecture Notes in Computer Science* 4819.

Yuan, Y. & Huang, T. (2005) A Matrix Algorithm for Mining Association Rules. *Lecture Notes in Computer Science* 3644.