JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Synthesis and Application of Global Negative Association Rules in Multi-Database Mining

[1]Sasikumar P, [2]K.T. Meena Abarna
[1]Research Scholar, [2]Associate Professor
Department of Computer Science and Engineering
Annamalai University, Annamalai Nagar, India

Abstract: Association rule mining has garnered widespread adoption within the data mining community, facilitating the discovery of relationships among frequently co-occurring item-sets. Alongside the exploration of positive association rules, the mining of negative association rules, aimed at identifying negation relationships among frequent item-sets, assumes equal significance. Notably, negative association rule mining plays a crucial role in customer-driven domains, such as market basket analysis, where the identification of conflicting products holds practical value. In the context of multi-database mining, the extraction of negation relationships among item-sets and the synthesis of global negative association rules from multiple geographically dispersed data sources become pivotal in shaping strategic and branch-level decisions. This research endeavor aims to propose a method for synthesizing global negative association rules, which are collectively endorsed by a majority of participating data sources during the mining process across diverse data repositories. To validate the theoretical foundation of the proposed approach, experimental data from the UCI machine learning repository data set are employed for rigorous testing. The investigation includes an in-depth analysis of space and time complexity, demonstrating the efficiency and efficacy of the proposed methodology.

Index Terms **- Negation relation, multi-databases, local pattern analysis, rule synthesizing.**

I. INTRODUCTION

Owing to the rapid advancements in communication network technologies and data dissemination, enterprise business organizations have expanded their branches across diverse geographical locations. This decentralized distribution of branches necessitates the efficient analysis and mining of branch data. Conventionally, the Mono-database mining technique has been employed to address such scenarios, where branch databases are integrated into a large repository known as a data warehouse for subsequent mining. However, this approach is plagued by several limitations, including information loss and privacy concerns pertaining to branch databases, the need for data movement from branch databases to a centralized warehouse, and significant investments required for deploying the necessary software and hardware.

To overcome these limitations, the adoption of the local pattern analysis strategy for distributed data mining, as an alternative to Mono-mining, has proven advantageous. In this approach, patterns mined from individual databases are forwarded instead of transmitting the entire databases, thereby mitigating issues related to heterogeneity of data integration and data ownership loss.

Association rule mining, a prominent research area in the Knowledge Discovery Process (KDD), holds significant potential in capturing customers' buying behavior. Besides its application in market basket analysis, association rule mining serves as the foundation for decision-making activities such as formulating promotional strategies, pricing, or product placements. An association rule is represented in the form of $X \rightarrow Y$, describing the relationship between item-sets X and Y, where X, $Y \subset I$, and $X \cap Y = \emptyset$. Each association rule is characterized by two quality measurements, namely Support and confidence [3], which assess its interestingness. Association rules satisfying minimum support and minimum confidence thresholds are considered strong rules, identifying positive relationships or co-occurrences of frequent item-sets.

In addition to positive association rules, the investigation of negation relations, which identify occurrences of one frequent item by the absence of others, is also crucial in data analysis and decision-making. If a positive association rule is represented as an implication $A \rightarrow B$, then negative association rules can be represented in one of the following forms: $A \rightarrow \neg B$, $\neg A \rightarrow B$, $\neg A \rightarrow \neg B$, where 'A' and 'B' are the frequent item-sets. Given that databases are distributed across different regions with distinct features, identifying negative relationships among forwarded patterns, pruning uninteresting rules, and synthesizing global negative association rules will be instrumental in formulating organizational policies.

1.1. Contributions

The principal contributions of this study are outlined below:

1. We adopt a local pattern analysis strategy to effectively mine negative associations among frequent item sets and subsequently synthesize global negative association rules based on the forwarded negation relations.
2. To synthesize global negative association rules extracted from multiple data sources, we introduce a transactions-population based weighting model. This model aids in the pruning of uninterested negative rules by utilizing the gateway threshold, referred to as the effective vote rate. Moreover, we employ site weights derived from the transactions-population of respective data sources to synthesize forwarded negative relations. Notably, this work represents a pioneering endeavor in the context of negative association rules within multi-database mining.
3. An in-depth analysis of the space and time complexity of the proposed approach for generating synthesized global association rules is conducted.

The subsequent sections of this paper are organized as follows:
- Section 2 elucidates the fundamental concepts and terminologies pertinent to the proposed approach.
- In Section 3, we expound upon the procedure employed to synthesize global negative association rules.
- Section 4 is devoted to experimental investigations, as well as a comprehensive analysis of space and time complexity.
- Section 5 presents a review of related research in the domains of negative association rule mining and synthesizing techniques employed in multi-database mining.
- Finally, Section 6 concludes the paper, summarizing the key findings and contributions.

II. Basic Concepts and Terminologies

This section elucidates the fundamental principles of negative association rule mining and introduces the terminologies that are integral to the proposed approach.

2.1. Negative Association Rules

Let A and B be frequent item sets, wherein the positive association rule is denoted as A→B. Corresponding negative association rules can be represented in various forms, such as (A→¬B), (¬B→A), or (¬A→¬B), where "¬A" and "¬B" signify the negation of item sets A and B, respectively. The symbol "¬" indicates the absence or negative attribute of an itemset. To generate robust negative association rules, the domain expert specifies minimum support and minimum confidence thresholds. The calculation of support and confidence for negative association rules is carried out as follows:

$$supp(\neg A) = 1 - supp(A)$$
$$supp(A \rightarrow \neg B) = supp(A) - supp(A \cup B)$$
$$supp(\neg A \rightarrow B) = supp(B) - supp(A \cup B)$$
$$supp(\neg A \rightarrow \neg B) = 1 - supp(A) - supp(B) + supp(A \cup B)$$
$$conf(A \rightarrow \neg B) = (supp(A) - supp(A \cup B)) / supp(A)$$
$$conf(\neg A \rightarrow B) = (supp(B) - supp(A \cup B)) / (1 - supp(A))$$
$$conf(\neg A \rightarrow \neg B) = (1 - supp(A) - supp(B) + supp(A \cup B)) / (1 - supp(A))$$

2.2. Site Weight of Participating Data Sources

Consider a scenario where we have a set of branch databases denoted as S1, S2, ..., Sm, actively engaged in the process of mining negative association rules. Let w's1, w's2, ..., w'sm represent the weights corresponding to the transaction populations of the individual data sources. Generally, w'sj denotes the un-normalized site weight of site j, calculated based on its respective transactions-population. The normalized weight of a site is determined by the ratio between the transactions-population within the specific data-source and the total of the transactions-populations of all participating sites.

$$\text{Normalized weight of Site } j = w_{sj} = \frac{w'sj}{\sum_{j=1}^{m} w'sj} \qquad (1)$$

2.3. Effective Vote Rate

In order to eliminate uninteresting rules that are forwarded from the local databases, a rule selection measure known as "min. γeffective" is employed as a threshold. For each negative association rule Ri, the value of γeffective is computed using the following calculation:

$$\gamma_{effective}(Ri) = \sum_{j=1}^{m} \delta(i, j) * wsj \qquad (2)$$

Let δ(i, j) be defined as follows: δ(i, j) = 1 if the rule Ri is present in site j; otherwise, δ(i, j) = 0. This function represents the percentage of votes received from various data sources for a given rule, determined based on the transaction populations of the corresponding data sources.

2.4. Synthesized Global Support and Confident of Negative Rules

The local support of a negative association rule Ri at site j is denoted as Suppj (Ri). Correspondingly, the local confidence of rule Ri at site j is represented as Confj (Ri). The support for the antecedent of the negative rule Ri at site j is denoted as Supp_antej (Ri). The computation of the synthesized global support and global confidence for the negative association rule Ri is carried out as follows:

$$Supp_G\,(Ri) = \sum_{j=1}^{m} wsj * Suppj(Ri) \qquad (3)$$

$$Conf_G(Ri) = \frac{Supp_G\,(R_i)}{Supp\_ante_G\,(R_i)} \qquad (4)$$

III. Proposed Procedure

The process of synthesizing global negative association rules entails several steps. Initially, a set of negative association rules is generated from the branch data sources, adhering to local support and local confidence thresholds. Subsequently, uninteresting negative association rules are eliminated using a rule selection threshold-effective vote rate. The global synthesis of negative association rules is then performed, considering site weights, local support, and local confidence values provided by individual sites. The globally synthesized negative association rules are those whose global support and global confidence values exceed the user-specified threshold.

Consider four sites, denoted as S1, S2, S3, and S4, each having respective transaction populations of 20,000, 10,000, 5,000, and 5,000. Employing a minimum support threshold of 0.25 and a minimum confidence threshold of 0.40, the four rules (R1, R2, R3, and R4) are voted by the four sites (S1, S2, S3, and S4), and their tabulated results can be observed in Table 1.

Table 1. Rules voted by the sites.

Rule	S1 (20000)	S2 (10000)	S3 (5000)	S4 (5000)
R1=A→¬B	0.45,0.60	0.27,0.40	0.35,0.50	0.30,0.60
R2= ¬ C→D	0.25,0.40	0.35,0.60	0.25,0.40	-
R3= ¬ E→¬F	0.40,0.50	-	-	-
R4= ¬ G→H	-	-	-	0.35,0.45

Weight of all data sources 'W'_S' is calculated as:

$$W'_S = 20000 + 10000 + 5000 + 5000 = 40000 \qquad (5)$$

Normalized site weights 'w_{si}' are calculated as:

$$w_{s1} = 0.50;\ w_{s2} = 0.25;\ w_{s3} = 0.125;\ w_{s4} = 0.125 \qquad (6)$$

Rule weights 'W_{Ri}' are calculated as:

$$\begin{aligned} W_{R1} &= 1 * 0.50 + 1 * 0.25 + 1 * 0.125 + 1 * 0.125 = 1.0 \\ W_{R2} &= 1 * 0.50 + 1 * 0.25 + 1 * 0.125 = 0.875 \\ W_{R3} &= 1 * 0.50 = 0.50 \\ W_{R4} &= 1 * 0.125 = 0.125 \end{aligned} \qquad (7)$$

By setting the Rule selection threshold at 0.50, rule R4 is excluded from the synthesis process. Despite both rule R3 and R4 being endorsed by a single site, rule R3 is retained in the synthesis process owing to its association with the transaction's population of data source S1. This highlights the significance of the rule selection measure, namely the effective vote rate. The ensuing global support and confidence values of the rules are subsequently computed and tabulated in Table 2.

IV. Experimental Results and Complexity Analysis

To assess the theoretical validity of the proposed approach, a series of experimental investigations was conducted using diverse datasets. This section elaborates on the experimental study performed on the mushroom dataset obtained from the UCI Machine Learning Repository. The dataset comprises 8,124 transactions with 120 distinct items, and the average transaction length is 24. For the analysis, the entire database was partitioned into four subsets, denoted as S1, S2, S3, and S4, with transaction populations of 3000, 2000, 2000, and 1124, respectively.

To generate negative frequent items and association rules from individual sites, a minimum local support threshold of 0.30 was chosen. Similarly, a global threshold level of 0.30 was considered for synthesizing rules forwarded from individual sites. The effective vote-rate was set at 0.66 to identify candidates for the synthesis process. Based on the respective transaction populations, the normalized site weight for S1, S2, S3, and S4 were calculated as 0.369276, 0.246184, 0.246184, and 0.138355, respectively.

Through this analysis, 24 negative association rules were identified as candidates for synthesis, supported by all four sites with an effective vote-rate of 1 (refer to Table 3). These rules are deemed strong negative association rules since they receive support from all participating data sources.

Additionally, 37 negative association rules were supported by three sites, meeting the effective vote-rate of 0.66, and were considered as candidates for the global rule synthesis process (refer to Table 4).

Table 2. Global negative association rule with synthesized support and confidence.

Rule	Synthesized Support	Synthesized Confidence
R1=A→¬B	0.37375	0.69375
R2= ¬ C→D	0.21875	0.536425
R3= ¬ E→¬F	0.20	0.40

$Supp_G\ (A \rightarrow \neg B)$ *is calculated as:*

$$\begin{aligned} &= (0.50*0.45) + (0.25*0.27) + (0.125*0.35) + (0.125*0.30) \\ &= 0.225 + 0.0675 + 0.04375 + 0.0375 \\ &= 0.37375 \end{aligned} \tag{8}$$

$Conf_G\ (A \rightarrow \neg B)$ *is calculated as:*

$$\begin{aligned} &= (0.50*(0.45/0.60)) + (0.25*(0.27/0.40)) + \\ &0.125*(0.35/0.50)) + (0.125*(0.30/0.60)) \\ &= 0.375 + 0.16875 + 0.0875 + 0.0625 \\ &= 0.69375 \end{aligned} \tag{9}$$

$Supp_G\ (\neg C \rightarrow D)$ *is calculated as:*

$$\begin{aligned} &= (0.50*0.25) + (0.25*0.35) + (0.125*0.25) \\ &= 0.10 + 0.0875 + 0.03125 \\ &= 0.21875 \end{aligned} \tag{10}$$

$Conf_G$ ($\neg C \rightarrow D$) is calculated as:

$$= (0.50*(0.25/0.40)) + (0.25*(0.35/0.60)) + (0.125*(0.25/0.40)) = 0.536425 \quad (11)$$

$Supp_G$ ($\neg E \rightarrow \neg F$) is calculated as:

$$= (0.50*0.40) = 0.20 \quad (12)$$

$Conf_G$($\neg E \rightarrow \neg F$) is calculated as:

$$= 0.50*(0.40/0.50) = 0.40 \quad (13)$$

Furthermore, 79 negative association rules were supported by only two sites and did not qualify as candidates for rule synthesis, as they failed to meet the minimum effective vote-rate threshold of 0.66. As a result, a total of 61 rules found in Table 2 and Table 3 were recognized as candidate rules. During the synthesis process, 25 negative association rules among the candidates failed to attain the global threshold value of 0.30. The remaining 36 rules were successfully synthesized as global negative association rules, and their corresponding synthesized support and confidence values are presented in Table 5.

4.1. Space and Time Complexity Analysis

This section presents a comprehensive analysis of the space and time complexity pertaining to the proposed approach.

Table 3. Candidature for global negative rule synthesizing voted by all the sites.

Rule-Id	Rule Name	Effective Vote Rate
1	¬ 36→ 34	1
2	¬ 10→ 36	1
3	¬ 63→ 90	1
4	¬ 59→ 90	1
5	¬ 53→ 90	1
6	¬ 53→ 86	1
7	¬ 63→ 86	1
8	¬ 63→ 85	1
9	¬ 59→ 86	1
10	¬ 59→ 85	1
11	¬ 53→ 85	1
12	¬ 63→¬ 59	1
13	¬ 36→ 90	1
14	¬ 36→ 86	1
15	¬ 36→ 85	1
16	¬ 10→ 86	1
17	¬ 10→ 85	1
18	36→¬ 63	1
19	36→ ¬ 59	1
20	¬ 53→ 36	1
21	¬ 52→ 85	1
22	¬ 63→34	1
23	¬ 59→34	1
24	¬ 53→34	1

Table 4. Candidature for global negative rule synthesizing and not supported by all sites.

Rule-Id	Rule Name	Supporting Sites	Effective Vote Rate
25	¬ 39→ 34	S1, S2, S3	0.86
26	¬ 93→39	S1,S2,S4	0.75
27	¬ 3→34	S1, S2, S3	0.86
28	¬ 28→¬ 93	S1,S2,S4	0.75
29	¬ 39→¬ 2	S1,S2,S4	0.75
30	¬ 6→34	S1, S2, S3	0.86
31	¬ 56→86	S1, S2, S3	0.86
32	¬ 93→¬ 2	S1,S2,S4	0.75
33	¬ 10→34	S1, S2, S3	0.86
34	¬ 28→¬ 2	S1,S2,S4	0.75
35	¬ 56→85	S1, S2, S3	0.86
36	¬ 39→86	S1, S2, S3	0.86
37	¬ 39→85	S1,S2,S4	0.75
38	¬ 6→90	S1, S2, S3	0.86
39	¬ 76→86	S1,S3,S4	0.75
40	¬ 53→¬ 63	S1, S2, S3	0.86
41	¬ 76→85	S1,S3,S4	0.75
42	¬ 6→86	S1, S2, S3	0.86
43	¬ 53→¬ 59	S1, S2, S3	0.86
44	¬ 10→90	S1, S2, S3	0.86
45	¬ 28→85	S1,S2,S4	0.75
46	¬ 6→85	S1, S2, S3	0.86
47	¬ 2→85	S1,S2,S4	0.75
48	¬ 3→90	S1, S2, S3	0.86
49	¬ 3→86	S1, S2, S3	0.86
50	¬ 3→85	S1, S2, S3	0.86
51	¬ 28→39	S1,S2,S4	0.75
52	¬ 76→34	S1,S3,S4	0.75
53	¬ 13→85	S1,S3,S4	0.75
54	¬ 52→86	S1, S2, S3	0.86
55	¬ 67→34	S1,S3,S4	0.75
56	¬ 67→86	S1,S3,S4	0.75
57	¬ 67→85	S1,S3,S4	0.75
58	¬ 56→34	S1, S2, S3	0.86
59	¬ 52→34	S1, S2, S3	0.86
60	¬ 93→85	S1,S2,S4	0.75
61	¬ 76→¬ 67	S1,S3,S4	0.75

Table 5. Global negative association rules with their synthesized support and confidence.

Rule-Id	Rule Name	Synthesized Support	Synthesized Confidence
2	¬ 10→36	0.4436	0.7385
3	¬ 63→90	0.3683	0.9386
4	¬ 59→90	0.3388	0.9335
5	¬ 53→90	0.3545	0.8191
6	¬ 53→86	0.4082	0.9431
7	¬ 63→86	0.3915	0.9975
8	¬ 63→85	0.3924	1.0000
9	¬ 59→86	0.3619	0.9973
10	¬ 59→85	0.3629	1.0000
11	¬ 53→85	0.4328	1.0000
16	¬ 10→86	0.5771	0.9607
17	¬ 10→85	0.6007	1.0000
18	36→¬ 63	0.3235	0.3858
20	¬ 53→36	0.3776	0.8726
21	¬ 52→85	0.5672	0.9584
22	¬ 63→34	0.3902	0.9944
23	¬ 59→34	0.3607	0.9940
24	¬ 53→34	0.4069	0.9403
27	¬ 3→34	0.4365	0.7573
30	¬ 6→34	0.4919	0.7780
31	¬ 56→86	0.3998	0.7400
33	¬ 10→34	0.5037	0.8385
35	¬ 56→85	0.4019	0.7439
38	¬ 6→90	0.4674	0.7393
42	¬ 6→86	0.4919	0.7780
44	¬ 10→90	0.4839	0.8055
45	¬ 28→85	0.3689	0.5998
46	¬ 6→85	0.4938	0.7811
48	¬ 3→90	0.4090	0.7097
49	¬ 3→86	0.4366	0.7575
50	¬ 3→85	0.4379	0.7599
53	¬ 13→85	0.4964	0.6685
54	¬ 52→86	0.4906	0.8650
58	¬ 56→34	0.3997	0.7398
59	¬ 52→34	0.4906	0.8650
60	¬ 93→85	0.3126	0.5594

4.1.1. Rule Selection

In this study, we consider a set of N' rules originating from individual sites denoted as Sj (j = 1, ..., m). Our objective is to utilize these rules as inputs for analysis. To store all the rules from the sites, a total of |Ri(S1)| + |Ri(S2)| + ... + |Ri(Sm)| units are required, where m represents the number of sites, and n is the maximum value among |Ri(S1)|, |Ri(S2)|, ..., |Ri(Sm)|. The number of rules in the rule set |Ri(Sj)| of site Sj is represented by |Ri(Sj)|. Additionally, N units are needed to store all the rules in the reduced set of rules, denoted as S.

Consequently, the space complexity can be expressed as O(N + mn). Moving on to the time complexity, we elucidate the steps involved: The N rules in the rule set S must be computed from the rule sets of the given sites. Each rule necessitates access to m sites. Accessing a site demands, at most, n comparisons, where n is the maximum value among |Ri(S1)|, |Ri(S2)|, ..., |Ri(Sm)| and |Ri(Sj)|. As a result, the time complexity is represented as O(Nnm).

4.1.2. Rule Synthesizing

In the process of synthesis, Global negative association rules are derived from the provided rule set S, wherein the space complexity of rule selection is denoted by O (N + mn). Here, m units are essential to store the weights corresponding to m sites, while N units are necessary to store the weights of all rules within S, where N ≤ N'. The reduced rule set S contains N' rules. To accommodate the support and confidence values for all rules in S, an additional 3N units are required. Consequently, the overall space complexity for synthesizing global negative association rules is expressed as O (N + mn + m + N + 3N) = O (N + mn). Likewise, the time complexity is calculated as O(Nnm).

V. Related Research Works

Brin et al. [5] introduced the notion of negative relationships between variables employing a chi-square based model. They further proposed a correlation metric to ascertain the nature of these relationships. Savasere et al. [10] explored the concept of negative association rule mining by combining positive frequent item-sets with domain knowledge represented as a taxonomy. Wu et al. [12] devised an algorithm capable of generating both positive and negative association rules, introducing a novel measure called mininterest to enhance frequent itemset pruning.

Antonie and Zaine [4] extended the support-confidence framework through the incorporation of a sliding correlation coefficient threshold, enabling the discovery of negative association rules exhibiting strong negative correlations between antecedents and consequents. Dong et al. [7] presented the Multiple Level Minimum Supports (MLMS) model, which concurrently explores infrequent and frequent item sets using multiple level minimum supports. Additionally, they proposed the Valid Association Rule Based on

Correlation Coefficient and Confidence (VARCC) measure [6], effectively generating positive and negative association rules from the frequent and infrequent item sets discovered by the MLMS model.

Zhang et al. [14] introduced the local pattern analysis strategy, mitigating limitations arising from centralized mining. Within this context, they introduced new pattern types such as high vote patterns, exceptional patterns, and suggested patterns. Synthesizing meaningful patterns from numerous local patterns forwarded by individual databases necessitates a synthesis process. Wu and Zhang [11] proposed a weighting model for synthesizing high-frequency rules from multiple databases. This model calculates the weight of a branch database based on the number of high-frequency rules supported by it.

Moreover, Zhang et al. [13] advocated an approach for synthesizing global exceptional patterns in multi-database mining applications. Adhikari et al. [2] introduced a pipelined feedback technique for mining multiple databases and conducted experimental investigations on various datasets. Building upon the Wu-Zhang model, Adhikari and Rao [1] attempted to synthesize heavy association rules in a multi-database environment and determined whether a heavy association rule is of high-frequent or exceptional nature. Ramkumar and Srinivasan [8] proposed a weighting model for synthesizing high-frequent association rules from multiple data sources based on the transactions-population of data sources. Their synthesized results closely aligned with centralized mining outcomes. Furthermore, they proposed a multi-level rule synthesis model for categorizing the synthesized patterns into various categories like sub-global patterns and local patterns [9].

VI. Conclusion

This research article presents a systematic approach for the synthesis of global negative association rules derived from multiple data sources. The proposed procedure serves as a valuable tool for discovering patterns characterized by low frequency and high correlation within the context of multi-database mining environments. The resultant global negative association rules play a crucial role in identifying item-sets that exhibit conflicts on a global scale, enabling organizational leaders to explore and capitalize on business opportunities associated with infrequent yet meaningful item-sets of interest.

VII. References

[1] Adhikari A. and Rao R., “Synthesizing Heavy Association Rules from Different Real Data Sources,” Pattern Recognition Letters, vol. 29, no. 1, pp. 59-71, 2008.

[2] Adhikari A., Rao R., Prasad B., and Adhikari J., “Mining Multiple Large Data Sources,” the International Arab Journal of Information Technology, vol. 7, no. 3, pp. 241-249, 2010.

[3] Agrawal R. and Srikant R., “Fast Algorithms for Mining Association Rules,” in Proceedings of the 20th International Conference on Very Large Databases, Santiago, Chile, pp. 478-499, 1994.

[4] Antonie M. and Zaïane R., “Mining Positive and Negative Association Rules: An Approach for Confined Rules,” in Proceedings of the 8th European Conference on Principles and Practice of Knowledge Discovery in Databases, New York, USA, pp. 27-38, 2004.

[5] Brin S., Motwani R., and Silverstein C., “Beyond Market Basket: Generalizing Association Rules to Correlations,” in Proceedings of ACM SIGMOD International Conference on Management of Data, New York, USA, pp. 265-276, 1997.

[6] Dong X., Niu Z., Shi X., Zhang X., and Zhu D., “Mining both Positive and Negative Association Rules from Frequent and Infrequent Itemsets,” in Proceedings of the 3rd International Conference on Advanced Data Mining and Applications, Harbin, China, vol. 4632, pp. 122-133, 2007.

[7] Dong X., Niu Z., Zhu D., Zheng Z., and Jia Q., “Mining Interesting Infrequent and Frequent Item sets Based on MLMS Model,” in Proceedings of the 4th International Conference on Advanced Data Mining and Applications, Chengdu, China, vol. 5139, pp. 444-451, 2008.

[8] Ramkumar T. and Srinivasan R., “Modified Algorithms for Synthesizing High-Frequency Rules from Different Data Sources,” Knowledge and Information System, vol. 17, no. 3, pp. 313-334, 2008.

[9] Ramkumar T. and Srinivasan R., “Multi-Level Synthesis of Frequent Rules from Different Data Sources,” the International Journal of Computer Theory and Engineering, vol. 2, no. 2, pp. 195-204, 2010.

[10] Savasere A., Omiecinski E., and Navathe S., “Mining for Strong Negative Associations in a Database of Customer Transactions,” in Proceedings of the 14th International Conference on Data Engineering, Orlando, USA, pp. 494-502,1998.

[11] Wu X. and Zhang S., “Synthesizing High-Frequency Rules from Different Data Sources,” IEEE Transactions on Knowledge and Data Engineering, vol. 15, no. 2, pp. 353-367, 2003.

[12] Wu X., Zhang C., and Zhang S., “Efficient Mining of both Positive and Negative Association Rules,” ACM Transactions on Information Systems, vol. 22, no. 3, pp. 381-405, 2004.

[13] Zhang C., Liu M., and Nie W., “Identifying Global Exceptional Patterns in Multi-Database Mining,” IEEE Computational Intelligence Bulletin, vol. 3, no. 1, pp. 19-24, 2004.

[14] Zhang S., Wu X., and Zhang C., “Multi-Database Mining,” IEEE Computational Intelligence Bulletin, vol. 2, no.1, pp. 5-13, 2003.